

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 1 043 399 A2

(12)

EUROPÄISCHE PATENTANMELDUNG

(43) Veröffentlichungstag:
11.10.2000 Patentblatt 2000/41

(21) Anmeldenummer: 00105929.4

(22) Anmeldetag: 23.03.2000

(51) Int. Cl.⁷: **C12N 15/86**, C12N 7/01,
C12N 7/04, C12N 5/10,
C07K 14/18, A61K 49/00,
A61K 48/00

(84) Benannte Vertragsstaaten:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE**
Benannte Erstrückungsstaaten:
AL LT LV MK RO SI

(30) Priorität: 03.04.1999 DE 19915178

(71) Anmelder:
**Bartenschlager, Ralf, Dr.
55239 Gau-Odernheim (DE)**

(72) Erfinder:
**Bartenschlager, Ralf, Dr.
55239 Gau-Odernheim (DE)**

(74) Vertreter: **Rudolph, Ulrike, Dr.
Patentanwältin
In der Schanz 10
69198 Schriesheim (DE)**

(54) **Hepatitis C Virus Zellkultursystem**

(57) Das erfindungsgemäße Hepatitis C Virus (HCV) Zellkultursystem besteht aus humanen Hepato-mazellen, die mit einem HCV-RNA-Konstrukt transfiziert sind, das die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zudem wenigstens ein selektierbares Markergen (Selektionsgen) umfaßt.

EP 1 043 399 A2

Beschreibung

[0001] Die Erfindung betrifft ein Hepatitis C Virus (HCV) Zellkultursystem, das im wesentlichen eukaryontische Zellen umfaßt, die eingeschleustes HCV-spezifisches Genmaterial enthalten, d.h. die mit HCV-spezifischem Genmaterial transfiziert sind.

[0002] Das Hepatitis C Virus (HCV) ist eine der Hauptursachen chronischer und sporadischer Leberkrankungen weltweit. Die meisten HCV-Infektionen verlaufen ohne erkennbare klinische Symptome, allerdings werden 80-90% der Infizierten dauerhafte Virusträger und bei 50% dieser dauerhaften Virusträger kommt es zu einer chronischen Leberentzündung mit unterschiedlichen Schweregraden. Ca. 20% der chronisch Infizierten entwickeln im Laufe von 10 bis 20 Jahren eine Leberzirrhose, auf deren Basis sich ein primäres Leberzellkarzinom entwickeln kann. Die chronische Hepatitis C ist heute die Hauptindikation für eine Lebertransplantation. Eine Kausaltherapie gibt es bisher noch nicht. Die einzige derzeit verfügbare Therapie ist die hochdosierte Verabreichung von Interferon-Alpha oder eine Kombination aus Interferon-Alpha und dem Purin-Nukleosidanalogon Ribavirin. Allerdings sprechen nur ca. 60 % aller Behandelten auf diese Therapie an und bei diesen kommt es in mehr als der Hälfte aller Fälle nach dem Absetzen der Behandlung zu einer erneuten Virämie.

Aufgrund der hohen Prävalenz, gerade auch in den Industrieländern, den schwerwiegenden Folgen chronischer Infektionen und dem Fehlen einer Kausaltherapie ist die Entwicklung einer HCV-spezifischen Chemotherapie ein wesentliches Ziel der pharmazeutischen Forschung und Entwicklung. Hauptproblem hierbei ist bisher das Fehlen eines geeigneten Zellkultursystems, das ein Studium der Virus-Replikation und der Pathogenese in eukaryontischen Zellen ermöglicht.

[0003] Aufgrund der geringen Virusmengen im Blut bzw. Gewebe, dem Fehlen geeigneter Zellkultursysteme oder Tiermodelle (bis heute ist der Schimpanse das einzige mögliche Versuchstier) sowie dem Fehlen effizienter Systeme zur Produktion virus-ähnlicher Partikel, konnte die molekulare Zusammensetzung des HCV-Partikels bis heute noch nicht eingehend untersucht bzw. aufgeklärt werden. Die derzeit vorliegenden Ergebnisse lassen sich wie folgt zusammenfassen: Das HCV ist ein umhülltes Plusstrang RNA Virus mit einem Partikeldurchmesser von 50-60 nm und einer mittleren Dichte von 1,03-1,1g/ml. Es wurde erstmals 1989 molekular kloniert und charakterisiert (Choo et al., 1989: *Science*, 244, 359-362). Die HCV-RNA hat eine Länge von ca. 9.6 kb (= 9600 Nukleotide), eine positive Polarität und besitzt ein einziges offenes Leseraster (ORF = open reading frame), das ein lineares Polyprotein von ca 3010 Aminosäuren kodiert (siehe Rice 1996, in *Virology*, B. N. Fields, D. M. Knipe, P. M. Howley, Eds. (Lippincott-Raven, Philadelphia, PA, 1996), vol. 1, pp.931-960; Clarke 1997, *J. Gen. Virol.* 78, 2397; und Bartenschlager 1997, *Intervirology* 40, 378 und vgl. Fig. 1 A). Bei der Virusreplikation wird das Polyprotein durch zelluläre und virale Proteasen in die reifen und funktionell aktiven Proteine gespalten.

Innerhalb des Polyproteins sind die Proteine wie folgt angeordnet (vom Amino- zum Carboxyterminus): Core-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B. Das Core-Protein ist die Hauptkomponente des Nukleokapsids. Die Glykoproteine E1 und E2 sind Transmembranproteine und Hauptkomponenten der Virushülle. Sie spielen wahrscheinlich bei der Anheftung des Virus an die Wirtszelle eine wesentliche Rolle. Diese drei Proteine Core, E1 und E2 bauen den Viruspartikel auf und werden deshalb als Strukturproteine bezeichnet. Die Funktion des Proteins p7 ist noch unklar. Das Protein NS2 ist wahrscheinlich die katalytische Domäne der NS2-3 Protease, die für die Prozessierung zwischen den Proteinen NS2 und NS3 verantwortlich ist. Das Protein NS3 hat zwei Funktionen, nämlich in der aminoterminalen Domäne eine Proteaseaktivität, die für die Polyproteinprozessierung essentiell ist, und in der carboxyterminalen Domäne eine NTPase/Helikase-Funktion, die wahrscheinlich bei der Replikation der viralen RNA eine Rolle spielt. Das Protein NS4A ist ein Kofaktor der NS3-Protease. Die Funktion des Proteins NS4B ist unbekannt.

[0004] Das offene Leseraster ist an seinem 5' Ende von einer ca 340 Nukleotide langen nicht-translatierten Region (NTR = non-translated region) flankiert, die als interne Ribosomenansatzstelle (IRES = internal ribosome entry site) fungiert, und an seinem 3' Ende von einer ca. 230 Nukleotide langen NTR, die höchstwahrscheinlich für die Genomreplikation von Bedeutung ist. Eine solche 3' NTR ist Gegenstand der Patentanmeldung PCT/US 96/14033. Die Strukturproteine in dem amino-terminalen Viertel des Polyproteins werden von der Signalpeptidase der Wirtszelle gespalten. Die Nicht-Strukturproteine (NS) 2 bis (NS) 5B werden von zwei viralen Enzymen prozessiert, nämlich von der NS2-3 und der NS3/4A Proteinase. Die NS3/4A Proteinase wird für alle Spaltungen jenseits des Carboxyterminus von NS3 benötigt. Die Rolle von NS4B ist nicht bekannt. NS5A, ein hoch phosphoryliertes Protein, scheint für die Interferon Resistenz verschiedener HCV-Genotypen verantwortlich zu sein (vgl. Enomoto et al. 1995, *J. Clin. Invest.* 96, 224; Enomoto et al. 1996, *N. Engl. J. Med.* 334, 77; Gale Jr. et al. 1997, *Virology* 230, 217; Kaneko et al. 1994, *Biochem. Biophys. Res. Commun.* 205, 320; Reed et al., 1997, *J. Virol.* 71, 7187) und NS5B wurde als die RNA-abhängige RNA Polymerase identifiziert.

[0005] Anhand dieser Erkenntnisse wurden erste Diagnosesysteme entwickelt, die entweder auf dem Nachweis von HCV-spezifischen Antikörpern in Patientenserum oder auf dem Nachweis von HCV-spezifischer RNA mittels RT-PCR (= Reverse Transcription Polymerase Chain Reaction) beruhen, und die mittlerweile routine- und/oder vorschrittmäßig bei allen Blutkonserven angewendet werden (müssen).

[0006] Seit der Erstbeschreibung des Genoms 1989 wurden mit Hilfe der PCR-Methode zahlreiche Teil- und Kompletsequenzen des HCV kloniert und charakterisiert. Ein Vergleich dieser Sequenzen zeigt eine hohe Variabilität des viralen Genoms, insbesondere im Bereich des NS5B-Gens, was letztendlich zu einer Einteilung in 6 Genotypen geführt hat, die selbst nochmals in Subtypen a, b, und c untergliedert sind. Die genomische Varianz ist nicht gleichmäßig über das Genom verteilt. So sind die 5'NTR und Teile der 3'NTR hoch konserviert, während bestimmte kodierende Sequenzen z.T. sehr stark variieren, vor allem die Hüllproteine E1 und E2.

[0007] Die klonierten und charakterisierten Teil- und Kompletsequenzen des HCV-Genoms wurden außerdem hinsichtlich geeigneter Angriffsziele für ein prospektives antivirales Therapeutikum untersucht. Dabei wurden drei virale Enzyme gefunden, die sich als solches Angriffsziel anbieten. Diese sind (1) der NS3/4A Proteasekomplex, (2) die NS3 Helikase und (3) die NS5B RNA-abhängige RNA Polymerase. Der NS3/4A Proteasekomplex und die NS3 Helikase konnten bereits kristallisiert und hinsichtlich ihrer dreidimensionalen Struktur aufgeklärt werden (Kim et al., 1996, Cell, 87, 343; Yem et al., 1998, Protein Science, 7, 837; Love et al., 1996, Cell, 87, 311; Kim et al., 1998, Structure, 6, 89; Yao et al., 1997, Nature Structural Biology, 4, 463; Cho et al., 1998, J. Biol. Chem., 273, 15045); bei der NS5B RNA-abhängigen RNA Polymerase ist dies bis heute noch nicht gelungen.

Obwohl mit diesen Enzymen bedeutsame Angriffsziele für eine Therapieentwicklung der chronischen HCV-Infektion definiert sind, und obwohl sowohl mit Hilfe von 'rational drug design' als auch mit Hilfe von 'high throughput screens' weltweit intensiv nach geeigneten Inhibitoren gesucht wird, leidet die Therapieentwicklung an einem großen Defizit, nämlich dem Fehlen von Zellkultursystemen oder einfachen Tiermodellen, die es erlauben, HCV-RNA oder HCV-Antigene direkt, zuverlässig und mit einfachen laborüblichen Methoden nachzuweisen. Das Fehlen solcher Zellkultursysteme ist auch der Hauptgrund dafür, daß das Verständnis der HCV-Replikation bis heute noch sehr lückenhaft und in weiten Teilen nur hypothetisch ist.

[0008] Obwohl nach Meinung der Fachwelt eine enge evolutionäre Beziehung zwischen HCV und den Flavi- und Pestiviren besteht und für diese autonom replizierende RNAs beschrieben sind, die in verschiedenen Zelllinien ohne weiteres zur Replikation gebracht werden können und dabei relativ hohe Ausbeuten zeigen (siehe Khromykh et al., 1997, J. Virol. 71, 1497; Behrens et al., 1998, J. Virol. 72, 2364; Moser et al., 1998, J. Virol. 72, 5318), waren ähnliche Versuche mit HCV bisher nicht erfolgreich.

[0009] Zwar ist aus verschiedenen Publikationen bekannt, daß Zelllinien oder primäre Zellkulturen mit HCV-haltigem, hochtitrigem Patientenserum infiziert werden können, (Lanford et al. 1994, Virology 202, 606; Shimizu et al. 1993, Proceedings of the National Academy of Sciences, USA, 90, 6037-6041; Mizutani et al. 1996, Journal of Virology, 70, 7219-7223; M. Ikeda et al. 1998, Virus Res. 56, 157; Fournier et al. 1998, J. Gen. Virol. 79, 2376 und darin zitierte Literaturstellen, Ito et al. 1998, Journal of General Virology, 77, 1043-1054), diese virusinfizierten Zelllinien oder Zellkulturen erlauben jedoch nicht den direkten Nachweis von HCV-RNA oder HCV-Antigenen. Die virale RNA in diesen Zellen ist weder in einem Northern-Blot (einem Standardverfahren zum quantitativen Nachweis von RNA) noch sind die viralen Proteine in einem Western-Blot oder mittels Immunpräzipitation detektierbar. Nur mit sehr aufwendigen und indirekten Methoden ist es überhaupt gelungen, eine HCV-Replikation nachzuweisen. Diese nachteiligen Umstände zeigen klar, daß die Replikation in diesen bekannten virusinfizierten Zelllinien oder Zellkulturen absolut unzureichend ist.

[0010] Desweiteren ist aus den Publikationen von Yoo et al. (1995, Journal of Virology, 69, 32-38) und von Dash et al., (1997, American Journal of Pathology, 151, 363-373) bekannt, daß Hepatomazelllinien mit synthetischer HCV-RNA, die mittels in vitro Transkription von kloniertem HCV-Genom gewonnen wurde, transfiziert werden können. In beiden Publikationen gingen die Autoren von dem Grundgedanken aus, daß das virale HCV-Genom eine Plusstrang-RNA ist, die nach dem Einschleusen in die Zelle direkt als mRNA fungiert, an die sich Ribosomen anheften und im Zuge von Translationsprozessen Virusproteine bilden, aus denen sich letztendlich neue HCV-Partikel bilden (können). Diese Virusreplikation, d.h. diese neu gebildeten HCV-Viren bzw. deren RNA wurde mittels RT-PCR nachgewiesen. Die publizierten Ergebnisse der durchgeführten RT-PCR sprechen jedoch dafür, daß die Effizienz der HCV-Replikation in den beschriebenen HCV-transfizierten Hepatomazellen nur sehr gering ist und jedenfalls nicht ausreicht, um Schwankungen in der Replikationsrate nach gezielter Einwirkung mit prospektiven antiviralen Therapeutika auch nur qualitativ, geschweige denn quantitativ zu messen. Außerdem ist im Stand der Technik mittlerweile bekannt (Yanagi et al., Proc. Natl. Acad. Sci. USA, 96, 2291-95, 1999), daß die hochkonservierte 3' NTR essentiell ist für die Virusreplikation, was in klarem Widerspruch zu den Behauptungen von Yoo et al. und Dash et al. steht, die für ihre Versuche in Unkenntnis des authentischen 3' Endes des HCV-Genoms ausschließlich HCV-Genome mit verkürzten 3' NTRs verwendet haben.

[0011] Aufgabe der vorliegenden Erfindung ist die Bereitstellung eines HCV- Zellkultursystems, bei dem die virale RNA in den transfizierten Zellen autonom und mit so hoher Effizienz repliziert, daß Schwankungen in der Replikationsrate nach gezielter Einwirkung mit virus- und insbesondere HCV-spezifischen prospektiven antiviralen Therapeutika qualitativ und quantitativ und mit Hilfe gängiger, laborüblicher Meßverfahren gemessen werden können.

[0012] Eine Lösung dieser Aufgabe besteht in der Bereitstellung eines Zellkultursystems der eingangs genannten Art, bei dem die eukaryontischen Zellen humane Zellen, insbesondere Hepatomazellen sind, die vorzugsweise von einer handelsüblichen Hepatomazelllinie abstammen, aber auch aus einer entsprechenden Primärzellkultur gewonnen sein können, und bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt ist, das im

wesentlichen die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR, vorzugsweise in der genannten Reihenfolge, und zudem wenigstens ein selektierbares Markergen (Selektionsgen) umfaßt. "NTR" steht hier und im folgenden für "nicht-translatierte Region" und ist dem einschlägigen Fachmann als Begriff bzw. Abkürzung bekannt und geläufig. Der Begriff "HCV-RNA-Konstrukt" umfaßt hier und im folgenden sowohl Konstrukte, die das komplette HCV-Genom enthalten, als auch solche, die lediglich einen Teil davon, d.h. ein HCV-Subgenom enthalten.

Eine bevorzugte Variante des erfindungsgemäßen Zellkultursystems, die sich in der Praxis sehr gut bewährt hat, ist unter der Nummer DSM ACC2394 (Laborbezeichnung HuBI 9-13) bei der DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH in Braunschweig, Deutschland, hinterlegt.

[0013] Mit dem erfindungsgemäßen Zellkultursystem wird erstmals ein in-vitro-System bereit gestellt, in dem HCV-RNA intrazellulär, autonom und in ausreichend großen Mengen repliziert und exprimiert wird, so daß eine quantitative Bestimmung sowohl der HCV-RNA-Mengen als auch der HCV-spezifischen Proteine mit konventionellen und zuverlässig genauen biochemischen Meßmethoden durchgeführt werden kann. Das heißt: es steht erstmals ein annähernd authentisches zellgestütztes ("cell-based") HCV-Replikationssystem zur Verfügung, das für die Entwicklung und Erprobung von antiviralen Pharmazeutika dringend benötigt wird. Dieses Testsystem bietet nun die Möglichkeit, potentielle Angriffsziele für eine wirksame HCV-spezifische Therapie zu identifizieren und HCV-spezifische Chemotherapeutika zu entwickeln und zu evaluieren.

[0014] Die Erfindung basiert auf der überraschenden Erkenntnis, daß eine effiziente Replikation der HCV-RNA nur dann in Zellen stattfindet, wenn diese mit einem HCV-RNA-Konstrukt transfiziert wurden, das mindestens die 5' und die 3' nicht-translatierten Regionen (NTR) und die Nichtstrukturproteine (NS) 3 bis 5B umfaßt und zusätzlich ein selektierbares Markergen (Selektionsgen) aufweist. Offensichtlich sind die Strukturgene für den Ablauf der Replikation ohne wesentliche Bedeutung, während andererseits eine effiziente Replikation der HCV-RNA anscheinend nur dann stattfindet, wenn die transfizierten Zellen einem permanenten Selektionsdruck unterzogen werden, der durch das mit der HCV-RNA verbundene selektierbare Markergen (Selektionsgen) vermittelt wird. Das Markergen (Selektionsgen) scheint somit einerseits die Selektion derjenigen Zellen zu provozieren, in denen die HCV-RNA produktiv repliziert, und andererseits scheint es die Effizienz der RNA-Replikation wesentlich zu steigern.

[0015] Gegenstand der Erfindung ist auch ein zellfreies HCV-RNA-Konstrukt, das sich dadurch auszeichnet, daß es die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR, vorzugsweise in der genannten Reihenfolge, und zudem ein selektierbares Markergen (Selektionsgen) umfaßt.

[0016] Der Begriff 5' NTR bzw. NS3 bzw. NS4A bzw. NS4B bzw. NS5A bzw. NS5B bzw. 3' NTR umfaßt im vorliegenden Zusammenhang jede Nukleotidsequenz, die im Stand der Technik als Nukleotidsequenz für den jeweils betreffenden funktionellen Abschnitt des HCV-Genoms beschrieben ist.

[0017] Die Bereitstellung eines solchen HCV-RNA-Konstrukts ermöglicht erstmals eine detaillierte Analyse der HCV - Replikation, - Pathogenese und - Evolution in Zellkulturen. Die HCV-spezifische virale RNA kann - als vollständiges Genom oder als Subgenom - gezielt in beliebigen Mengen erzeugt werden, und es besteht die Möglichkeit, das RNA-Konstrukt zu manipulieren und damit die HCV-Funktionen auf genetischer Ebene zu untersuchen und aufzuklären.

[0018] Da alle zur Zeit als Hauptangriffsziel für eine Therapie untersuchten HCV-Enzyme, nämlich die NS3/4A Protease, die NS3 Helikase und die NS5B Polymerase, in dem erfindungsgemäßen HCV-RNA-Konstrukt enthalten sind, kann es für alle entsprechenden Untersuchungen benutzt werden.

[0019] Eine Ausführungsform des HCV-RNA-Konstrukts, die sich in der praktischen Anwendung sehr gut bewährt hat, zeichnet sich dadurch aus, daß sie die Nukleotidsequenz gemäß Sequenzprotokoll SEQ ID NO:1 umfaßt. Weitere Ausführungsvarianten mit vergleichbar guten Eigenschaften für den Einsatz in der Praxis sind dadurch gekennzeichnet, daß sie eine Nukleotidsequenz entweder gemäß Sequenzprotokoll SEQ ID NO:2 oder SEQ ID NO:3 oder SEQ ID NO:4 oder SEQ ID NO:5 oder SEQ ID NO:6 oder SEQ ID NO:7 oder SEQ ID NO:8 oder SEQ ID NO:9 oder SEQ ID NO:10 oder SEQ ID NO:11 umfassen.

[0020] Es besteht die Möglichkeit, das erfindungsgemäße HCV-Subgenom-Konstrukt mit einer 3' NTR zu versehen, die eine im Stand der Technik hierfür bisher unbekannte Nukleotidsequenz aufweist, nämlich eine Nukleotidsequenz, die aus der Gruppe der nachfolgend aufgelisteten Nukleotidsequenzen (a) bis (i) ausgewählt ist:

- 5 (a) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTA
GCTTTTTTTTTTTCTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCC
AGACTGGAGT
- 10 (b) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAGTCT TTTTTTTTTTCT TTTTTTTTGA GAGAGAGAGT CTCACCTCTGT
TGCCAGACT GGAGC
- 15 (c) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTAATCTTT TTTTTTTTTCT TTTTTTTTGA GAGAGAGAGT CTCACCTCTGT
TGCCAGACT GCAGC
- 20 (d) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAGTCT TTTTTTTTTT TCTTTTTTTT TGAGAGAGAG AGTCTCACTC
TGTTGCCAG ACTGGAGT
- 25 (e) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAGTCT TTTTTTTTTT TCTTTTTTTT TGAGAGAGAG AGTCTCACTC
TGTTGCCAG ACTGGAGT
- 30 (f) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAGTCT TTTTTTTTTT TCTTTTTTTT TTGAGAGAGA GAGTCTCACT
CTGTTGCCCA GACTGGAGT
- 35 (g) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAGTCT TTTTTTTTTT CTTTTTTTTT GAGAGAGAGA
GTCTCACTCT GTTGCCAGACTGGAGT
- 40 (h) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTTAAT CTTTTTTTTT TTTTCTTTT TTTTGAGAGA
GAGAGTCTCA CTCTGTTGCC CAGACTGGAG T
- 45 (i) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTTAATC TTTTTTTTTT TTTTCTTTT TTTTGAGAG
AGAGAGTCTC ACTCTGTTGC CCAGACTGGA GT

50 Das in den erfindungsgemäßen HCV-RNA-Konstrukten enthaltene selektierbare Markergen (Selektionsgen) ist vorzugsweise ein Resistenzgen, insbesondere eine Antibiotikumresistenzgen.

Das hat den Vorteil, daß die mit diesem Konstrukt transfizierten Zellen leicht von den nicht transfizierten Zellen selektiert werden können, indem dem Zellkulturmedium z.B. im Fall eines Antibiotikumresistenzgens das betreffende Antibiotikum zugegeben wird. Unter 'Antibiotikum' wird im vorliegenden Zusammenhang jede Substanz verstanden, die die nicht-transfizierten Wirtszellen oder die Zellen, in denen die HCV-RNA nur mit geringer Effizienz repliziert, am Leben oder Wachstum hindert, insbesondere Zellgifte wie z.B. Puromycin, Hygromycin, Zeocin, Bleomycin oder Blasticidin.

55 [0021] Ein bevorzugtes selektierbares Markergen (Selektionsgen) bzw. Resistenzgen, das sich in der Praxis sehr gut bewährt hat, ist das Neomycinphosphotransferasegen.

[0022] Eine Alternative zu den Antibiotikumresistenzgenen ist z.B. das Thymidin-Kinase-Gen, mit dem eine HAT-

Selektion durchgeführt werden kann.

[0023] Die Position des selektierbaren Markergens (Selektionsgens), bzw. des bevorzugten Resistenzgens bzw. des besonders bevorzugten Antibiotikumresistenzgens in dem HCV-RNA-Konstrukt liegt vorzugsweise hinter der HCV 5' NTR, d.h. strangabwärts der 5' NTR bzw. strangaufwärts des HCV-Leserasters. Denkbar ist aber auch eine Insertion im Bereich der 3' NTR oder an anderer Stelle für ein HCV-Genom oder -Subgenom, z.B. innerhalb des Polyproteins.

[0024] Bei einer alternativen Ausführungsform des erfindungsgemäßen HCV-RNA-Konstrukts ist das selektierbare Markergen (Selektionsgen), insbesondere ein Antibiotikumresistenzgen, über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA bzw. der HCV-Genom- oder -Subgenomsequenz verbunden.

[0025] Damit geht der Vorteil einher, daß nach erfolgter Selektion derjenigen Zellen, in denen die HCV-RNA produktiv repliziert, in den daraus gewonnenen Zellklonen das Resistenzgen durch ribozymvermittelte Spaltung von der HCV-Subgenomsequenz abgetrennt werden kann, nämlich durch Aktivierung des einklonierten Ribozyms oder, im Fall eines Konstrukts mit einer Erkennungsstelle für ein Ribozym, durch Einschleusen des Ribozyms in die Zellen (z.B. mittels Transfektion eines Ribozymkonstrukts oder Infektion mit einem viralen Expressionsvektor, in den das entsprechende Ribozym eingesetzt wurde). Auf diese Weise wird ein authentisches HCV-Genom-Konstrukt ohne Resistenzgen erhalten, das zur Bildung authentischer infektiöser Viruspartikel befähigt ist.

[0026] Eine weitere bevorzugte Ausführungsform des erfindungsgemäßen HCV-RNA-Konstrukts zeichnet sich dadurch aus, daß das Konstrukt wenigstens ein integriertes Reportergen aufweist.

[0027] Unter Reportergen wird im folgenden jedes Gen verstanden, dessen Anwesenheit sich nach Überführung in einen Zielorganismus leicht und im allgemeinen mit einfachen biochemischen oder auch histochemischen Methoden nachweisen läßt, d.h. das für ein Protein kodiert, welches auch in geringen Mengen einfach und zuverlässig mit den laborüblichen Meßmethoden nachgewiesen und quantifiziert werden kann.

[0028] Diese Variante des HCV-RNA-Konstrukts hat den Vorteil, daß der Umfang der Replikation dieses Konstrukts anhand des Reportergenprodukts einfach und schnell mit laborüblichen Methoden gemessen werden kann.

[0029] Das Reportergen ist vorzugsweise ein Gen aus der Gruppe der Luziferasene, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), dem GFP-Gen (green-fluorescence-protein-Gen), dem GUS-Gen (Glukuronidasegen) oder dem SEAP-Gen (Sezernierte-Alkalische-Phosphatase-Gen). Diese Reportergene bzw. deren Produkte, nämlich die entsprechenden Reporterproteine, können z.B. mittels Fluoreszenz, Chemilumineszenz, colorimetrisch oder mit Hilfe immunologischer Methoden (z.B. ELISA) bestimmt werden.

[0030] Als Reportergen kommt aber auch ein Surrogatmarkergen in Betracht. Darunter sind in diesem Zusammenhang solche Gene zu verstehen, die für zelluläre Proteine, Nukleinsäuren oder — allgemein — für solche Funktionen kodieren, die einer von der Virusreplikation abhängigen Variation unterliegen, und die infolgedessen in denjenigen Zellen, in denen sich das HCV bzw. das HCV-RNA-Konstrukt vermehrt, entweder reprimiert oder aktiviert werden. Das heißt: die Reduktion bzw. Aktivierung dieser Funktion ist ein Ersatzmarker für die Virusreplikation bzw. die Replikation des HCV-RNA-Konstrukts.

[0031] Die Positionen von Reportergen und selektierbarem Markergen (Selektionsgen) können so gewählt sein, daß ein aus den beiden Genprodukten gebildetes Fusionsprotein exprimiert wird. Hierbei besteht die vorteilhafte Möglichkeit, daß diese beiden Gene so in dem HCV-RNA-Konstrukt angeordnet sind, daß ihre beiden exprimierten Proteine zunächst über eine Schnittstelle für eine Protease (z.B. Ubiquitin) oder über ein selbstspaltendes Peptid (z.B. das 2A-Protein der Picornaviren) fusioniert sind und erst später proteolytisch wieder getrennt werden.

Ebensogut können diese beiden Positionen aber auch derart getrennt voneinander liegen, daß beide Genprodukte separat exprimiert werden. (z.B. in der Reihenfolge: Marker- bzw. Resistenzgen — interne Ribosomenbindungsstelle — Reportergen).

Im Fall des Reportergens hat sich eine Ausführungsvariante besonders bewährt, bei der das Reportergen in das offene Leseraster des HCV-Genoms oder -Subgenoms einkloniert ist, und zwar derart, daß es erst nach einer proteolytischen Prozessierung in eine aktive Form überführt wird.

[0032] Das erfindungsgemäße Zellkultursystem in allen seinen Variationen kann für vielfältige Zwecke eingesetzt werden. Diese umfassen:

- Das Auffinden antiviral wirksamer Substanzen. Dies können beispielsweise sein: organische Verbindungen, die unmittelbar oder mittelbar in die Virusvermehrung eingreifen (z.B. Inhibitoren der viralen Proteasen, der NS3-Helikase, der NS5B RNA-abhängigen RNA Polymerase), antisense Oligonukleotide, die an eine beliebige Zielsequenz innerhalb des HCV-RNA-Konstrukts (z.B. die 5' NTR) hybridisieren und unmittelbar oder mittelbar zu einer Beeinflussung der Virusvermehrung führen z.B. auf Grund einer Reduktion der Translation des HCV-Polyproteins oder Ribozyme, die eine beliebige HCV-RNA-Sequenz spalten und damit die Virusreplikation beeinträchtigen.
- Die Evaluierung jeglicher Art antiviral wirksamer Substanzen in Zellkultur. Solche Substanzen können beispielsweise mittels 'rational drug design' oder 'high-throughput screening' am isolierten gereinigten Enzym gefunden werden. Unter Evaluierung sind vor allem die Bestimmung der inhibitorischen Eigenschaften der entsprechenden Substanz sowie deren Wirkungsmechanismus zu verstehen.

- Die Identifikation neuer Angriffsziele, viralen oder zellulären Ursprungs, für eine HCV-spezifische antivirale Therapie. Ist beispielsweise ein zelluläres Protein essentiell für die Virusreplikation, kann mittels Hemmung dieses zellulären Proteins die Virusreplikation ebenfalls beeinflusst werden. Das Auffinden solcher auxiliären Faktoren ist mit dem erfindungsgemäßen System ebenfalls möglich.
- 5 • Der Einsatz für die Resistenzbestimmung. Es ist anzunehmen, daß auf Grund der hohen Mutationsrate des HCV-Genoms Therapieresistenzen auftreten. Solche Resistenzen, die gerade bei der klinischen Zulassung einer Substanz von großer Bedeutung sind, lassen sich mit dem erfindungsgemäßen Zellkultursystem ermitteln. Zelllinien in denen sich das HCV-RNA-Konstrukt bzw. das HCV-Genom oder - Subgenom repliziert, werden mit steigenden Konzentrationen der entsprechenden Substanz inkubiert und die Replikation der viralen RNA wird entweder
10 anhand eines eingebrachten Reporters oder durch qualitative oder quantitative Bestimmung der viralen Nukleinsäuren oder Proteine bestimmt. Resistenz ist dann gegeben, wenn bei normaler Wirkstoffkonzentration keine Hemmung der Replikation zu beobachten ist. Durch Reklonierung der HCV-RNA (z.B. mittels RT-PCR) und Sequenzanalyse können die für Therapieresistenz verantwortlichen Nukleotid- bzw. Aminosäureaustausche ermittelt werden. Durch Einklonieren der/des entsprechenden Austausch/s in das Ursprungs-konstrukt kann deren
15 Kausalität für die Therapieresistenz bewiesen werden.
- Die Produktion von authentischen Virusproteinen (Antigene) für die Entwicklung und/oder Evaluierung von Diagnostika. Das erfindungsgemäße Zellkultursystem erlaubt auch die Expression von HCV-Antigenen in Zellkulturen. Diese Antigene können prinzipiell auch für den Aufbau diagnostischer Nachweisverfahren eingesetzt werden.
- Die Produktion von HCV Viren und virus-ähnlichen Partikeln insbesondere zur Entwicklung oder Herstellung von
20 Therapeutika und Impfstoffen sowie für diagnostische Zwecke. Insbesondere zellkultur-adaptierte vollständige HCV-Genome, die mit dem erfindungsgemäßen Zellkultursystem hergestellt werden können, sind in der Lage, mit hoher Effizienz in Zellkulturen zu replizieren. Diese Genome besitzen alle Funktionen des HCV und sind deshalb in der Lage infektiöse Viren zu produzieren.
- 25 **[0033]** Das erfindungsgemäße HCV-RNA-Konstrukt für sich genommen kann in allen seinen Variationen ebenfalls für vielfältige Zwecke eingesetzt werden. Dazu gehören vor allem:
 - Die Konstruktion attenuierter Hepatitis C Viren bzw. HCV-ähnlicher Partikel und deren Produktion in Zellkulturen: Durch zufällige oder gezielt hervorgerufene Mutationen, beispielsweise Punktmutationen, Deletionen oder Insertio-
30 nen, können attenuierte HCV- oder HCV-ähnliche Partikel erzeugt werden, d.h. Viren bzw. virusähnliche Partikel mit voller Replikationskompetenz aber verringerter bzw. fehlender Pathogenität. Solche attenuierte HCV- oder HCV-ähnliche Partikel sind insbesondere als Impfstoff einsetzbar.
 - Die Konstruktion von HCV-RNA-Konstrukten mit integrierten Fremdgenen, beispielsweise zur Verwendung als leberzellspezifische Genfährten in der Gentherapie. Auf Grund des ausgeprägten Leberzelltropismus des HCV und der Möglichkeit, Teile des Genoms durch heterologe Sequenzen zu ersetzen, lassen sich HCV-RNA-Konstrukte
35 herstellen, bei denen beispielsweise die Strukturproteine durch ein therapeutisch wirksames Gen ersetzt werden. Das so erhaltene HCV-RNA-Konstrukt wird in Zellen eingeschleust, vorzugsweise mittels Transfektion, die die fehlenden HCV-Funktionen, beispielsweise die Strukturproteine, konstitutiv oder induzierbar exprimieren. Durch diese dem Fachmann unter dem Begriff der 'Transkomplementation' bekannte Technik lassen sich Viruspartikel erzeugen, in die das HCV-RNA-Konstrukt eingebaut wird. Die so erhaltenen Partikel können für die Infektion vorzugsweise von Leberzellen verwendet werden. In diesen wird das therapeutisch wirksame Fremdgen zur Expression gebracht und entfaltet damit seine therapeutische Wirkung.
 - Das Auffinden permissiver Zellen, d.h. Zellen, in denen eine produktive Virusvermehrung erfolgt. Zu diesem Zweck wird entweder eines der vorgenannten HCV-RNA-Genomkonstrukte verwendet, das zur Bildung kompletter infektiöser Viren befähigt ist, oder es wird eines der vorgenannten HCV-Subgenom-Konstrukte eingesetzt, das aller-
45 dings zunächst gemäß vorgenanntem Beispiel in eine Zelllinie transfiziert wird, die die fehlenden Funktionen konstitutiv oder induzierbar exprimiert. In all diesen Fällen entstehen Viruspartikel, die zusätzlich zur HCV-Sequenz ein Resistenz- und/oder Reportergen tragen. Zum Auffinden von Zellen, in denen das HCV replizieren kann, werden diese Zellen mit den so hergestellten Viren infiziert und einer Antibiotikumselektion unterzogen oder, in Abhängigkeit vom HCV-RNA-Konstrukt, mittels Nachweis der Expression des Reportergens untersucht. Da eine
50 Antibiotikumresistenz bzw. eine Expression des Reportergens nur dann nachweisbar ist, wenn das HCV-RNA-Konstrukt repliziert, müssen die so gefundenen Zellen permissiv sein. Auf diese Weise lassen sich nahezu beliebige Zelllinien oder primäre Zellkulturen hinsichtlich der Permissivität testen und auffinden.
- 55 **[0034]** Das erfindungsgemäße Zellkultursystem erlaubt auch das gezielte Auffinden von HCV-RNA-Konstrukten, bei denen es auf Grund von Mutationen, die sich entweder zufällig im Rahmen der HCV-RNA-Replikation ereignen oder die gezielt in das Konstrukt eingeführt werden, zu einer Steigerung der Replikationseffizienz kommt. Solche Mutationen, die zu einer Veränderung der Replikation des HCV-RNA-Konstrukts führen, sind dem Fachmann als adaptive

Mutationen bekannt. Die Erfindung umfaßt deshalb auch Verfahren zur Gewinnung von zellkultur-adaptierten Mutanten eines erfindungsgemäßen HCV-RNA-Konstrukts gemäß vorstehender Beschreibung, wobei die Mutanten gegenüber dem originären HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen. Sie umfaßt desweiteren ein Verfahren zur Herstellung von Mutanten eines HCV-RNA-Vollängengenoms oder eines HCV-RNA-Teilgenoms oder eines beliebigen HCV-RNA-Konstrukts mit im Vergleich zu dem ursprünglichen HCV-RNA-Vollängengenom oder -Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz, sowie zellkultur-adaptierte Mutanten von HCV-RNA-Konstrukten, HCV-Vollängengenomen und HCV-Teilgenomen mit im Vergleich zu den ursprünglichen Konstrukten, Teil- oder Vollängengenomen erhöhter Replikationseffizienz.

[0035] Das erfindungsgemäße Verfahren zur Gewinnung von zellkultur-adaptierten Mutanten eines erfindungsgemäßen HCV-RNA-Konstrukts, wobei die Mutanten gegenüber dem HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen, ist dadurch gekennzeichnet, daß man ein Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt mit Selektionsgen nach einem der Ansprüche 4 bis 19 ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, daß man die gewachsenen Zellklone erntet, und daß man aus diesen Zellklonen die HCV-RNA-Konstrukte isoliert.

[0036] Bei einer vorteilhaften Weiterbildung dieses Herstellungsverfahrens werden die isolierten HCV-RNA-Konstrukte wenigstens einmal erneut passagiert, nämlich in Zellen eines Zellkultursystems nach Anspruch 1 eingeschleust, das dabei erhaltene Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial das isolierte HCV-RNA-Konstrukt mit Selektionsgen ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, die gewachsenen Zellklone geerntet und daraus die HCV-RNA-Konstrukte isoliert.

Mit dieser Verfahrensvariante kann der Grad der adaptiven Mutationen und damit der Grad der Replikationseffizienz in den betreffenden HCV-RNA-Konstrukten noch gesteigert werden.

[0037] Das erfindungsgemäße Verfahren zur Herstellung von Mutanten eines HCV - Vollängengenoms oder eines HCV- Teilgenoms oder eines beliebigen HCV-RNA-Konstrukts mit im Vergleich zu dem ursprünglichen HCV- Vollängengenom oder - Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz zeichnet sich dadurch aus, daß man mit Hilfe eines der beiden vorstehend genannten Herstellungsverfahren eine zellkultur-adaptierte Mutante eines HCV-RNA-Konstrukts herstellt, diese aus den Zellen isoliert, mit im Stand der Technik bekannten Methoden kloniert und sequenziert und durch Vergleich mit der Nukleotid- und Aminosäuresequenz des ursprünglichen HCV-RNA-Konstrukts die Art, Anzahl und Positionen der Mutationen bestimmt, und diese Mutationen dann entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, welche die betreffenden Mutationen enthalten, in ein (isoliertes) HCV-Vollängen- oder -teilgenom oder ein beliebiges HCV-RNA-Konstrukt einführt.

Zum Nachweis bzw. zur Verifizierung derjenigen Mutationen, die tatsächlich eine Veränderung der Replikation und insbesondere eine Replikationssteigerung bewirken, kann ein Test durchgeführt werden, bei dem die bestimmten Nukleotid- und/oder Aminosäureaustausche in das ursprüngliche HCV-RNA-Konstrukt eingeführt und dieses wiederum in Zellkultur eingeschleust wird. Wenn die eingeführte Mutation tatsächlich zu einer Steigerung der Replikation führt, sollte im Fall eines HCV-RNA-Konstrukts mit selektierbarem Markergen die Zahl der resistenten Zellklone bei dem künstlich mutierten Konstrukt deutlich höher sein als bei dem unbehandelten Konstrukt. Im Fall eines Konstrukts mit einem Reportergen sollte die Aktivität bzw. Menge des Reporters bei dem künstlich mutierten Konstrukt deutlich höher sein als bei dem unbehandelten.

[0038] Die erfindungsgemäßen zellkultur-adaptierten HCV-RNA-Konstrukte mit hoher Replikationseffizienz sind dadurch gekennzeichnet, daß sie durch Nukleotid- und/oder Aminosäureaustausche von einem HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 19 ableitbar sind und daß sie mit einem der beiden vorstehend genannten Herstellungsverfahren erhältlich sind.

[0039] Diese zellkultur-adaptierten HCV-RNA-Konstrukte können dazu verwendet werden, beliebige HCV-RNA-Konstrukte oder HCV-Vollängen- oder Teilgenome mit erhöhter Replikationseffizienz herzustellen. Dabei können sowohl Konstrukte mit einem selektierbaren Resistenzgen als auch Konstrukte ohne ein solches bzw. mit einem nicht-selektierbaren Reportergen (z.B. Luziferase) hergestellt werden, denn aufgrund der sehr hohen Replikationseffizienz des zellkultur-adaptierten HCV-RNA-Konstrukts kann dessen Replikation auch in nicht-selektierten Zellen nachgewiesen werden.

Die erfindungsgemäßen zellkultur-adaptierten Mutanten eines HCV-RNA-Konstrukts oder eines HCV-Vollängengenoms oder eines HCV-Teilgenoms mit im Vergleich zu dem ursprünglichen HCV-RNA-Konstrukt oder dem ursprünglichen HCV-Vollängengenom erhöhter Replikationseffizienz, sind dadurch charakterisiert, daß sie mit einem Verfahren erhältlich sind, bei dem man in einem zellkultur-adaptierten HCV-RNA-Konstrukt durch Sequenzanalyse und Sequenzvergleich die Art und Anzahl der Mutationen bestimmt und diese Mutationen in ein HCV-RNA-Konstrukt, insbesondere in ein HCV-RNA-Konstrukt gemäß einem der Ansprüche 4 bis 19, oder in ein (isoliertes) HCV-RNA-Vollängengenom einführt, entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten.

[0040] Eine Gruppe ganz bevorzugter HCV-RNA-Konstrukte, HCV-Vollängengenome und HCV-Teilgenomen mit hoher und sehr hoher Replikationseffizienz und infolgedessen sehr guter Eignung für die praktische Anwendung ist

EP 1 043 399 A2

dadurch gekennzeichnet, daß sie einen oder mehrere oder alle der in Tabelle 3 aufgelisteten Aminosäure- bzw. Nukleotidaustausche und/oder einen oder mehrere der folgenden Aminosäureaustausche aufweist: 1283 arg -> gly , 1383 glu -> ala , 1577 lys -> arg , 1609 lys -> glu , 1936 pro -> ser , 2163 glu -> gly , 2330 lys -> glu , 2442 ile -> val. (Die Zahlen beziehen sich auf die Aminosäurepositionen des Polyproteins des HCV-Isolats con1, siehe Tabelle 1).

5

Besondere Eigenschaften der in den Sequenzprotokollen angegebenen Sequenzen:

SEQ ID-NO: 1

10 **[0041]**

Name: I389/Core-3'/wt
Aufbau (Nukleotidpositionen):

15

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters

20

4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B
5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7

25

9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)

30

14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

SEQ ID-NO: 2

35 **[0042]**

Name: I337/NS2-3'/wt
Aufbau (Nukleotidpositionen):

40

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1181: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1190-1800: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters

45

4. 1801-8403: HCV Polyprotein von Nichtstrukturprotein 2 bis Nichtstrukturprotein 5B
5. 1801-2451: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
6. 2452-4344: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
7. 4345-4506: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
8. 4507-5289: Nichtstrukturprotein 4B (NS4B)
9. 5290-6630: Nichtstrukturprotein 5A (NS5A)
10. 6631-8403: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
11. 8407-8637: HCV 3' nicht-translatierte Region

55

EP 1 043 399 A2

SEQ ID-NO: 3

[0043]

- 5 Name: I389/NS3-3'/wt
Aufbau (Nukleotidpositionen):
1. 1-341: HCV 5' nicht-translatierte Region
 2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
 - 10 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
 4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B
-
5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
 - 15 6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
 7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
 8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
 9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
 10. 7771-8001: HCV 3' nicht-translatierte Region

SEQ ID-NO: 4

[0044]

- 25 Name: I337/NS3-3'/wt
Aufbau (Nukleotidpositionen):
1. 1-341: HCV 5' nicht-translatierte Region
 2. 342-1181: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
 - 30 3. 1190-1800: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
 4. 1801-7758: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B
-
5. 1801-3696: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
 - 35 6. 3697-3858: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
 7. 3859-4641: Nichtstrukturprotein 4B (NS4B)
 8. 4642-5982: Nichtstrukturprotein 5A (NS5A)
 9. 5983-7755: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
 10. 7759-7989: HCV 3' nicht-translatierte Region

SEQ ID-NO: 5

[0045]

- 45 Name: I389/NS2-3'/wt
Aufbau (Nukleotidpositionen):
1. 1-341: HCV 5' nicht-translatierte Region
 2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
 - 50 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
 4. 1813-8418: HCV Polyprotein von Nichtstrukturprotein 2 bis Nichtstrukturprotein 5B
-
5. 1813-2463: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
 - 55 6. 2464-4356: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
 7. 4357-4518: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
 8. 4519-5301: Nichtstrukturprotein 4B (NS4B)
 9. 5302-6642: Nichtstrukturprotein 5A (NS5A)

EP 1 043 399 A2

10. 6643-8415: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
11. 8419-8649: HCV 3' nicht-translatierte Region

SEQ ID-NO: 6

5

[0046]

Name: I389/NS3-3'/9-13F
Aufbau (Nukleotidpositionen):

10

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
- 15 4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 9-13F

15

20

5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

25

SEQ ID-NO: 7

[0047]

Name: I389/Core-3'/9-13F
Aufbau (Nukleotidpositionen):

30

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
- 35 4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 9-13F

35

40

45

5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7
9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

50

SEQ ID-NO: 8

[0048]

Name: I389/NS3-3'/5.1
Aufbau (Nukleotidpositionen):

55

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker

EP 1 043 399 A2

3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 5.1

5

5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5B (NS5A)

10

9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

SEQ ID-NO: 9

15 [0049]

Name: I389/Core-3'/5.1

Aufbau (Nukleotidpositionen):

20

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 5.1

25

5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7
9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

30

35

SEQ ID-NO: 10

40 [0050]

Name: I389/NS3-3'/19

Aufbau (Nukleotidpositionen):

45

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 19

50

5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

55

SEQ ID-NO: 11

[0051]

5 Name: I389/Core-3'/19
Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
- 10 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 19
5. 1813-2385: HCV Core Protein; Strukturprotein
- 15 6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7
9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
- 20 11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

25 [0052] Die Erfindung wird im folgenden anhand von Ausführungsbeispielen und dazugehörigen Tabellen und Figuren näher erläutert. Die erwähnten Figuren zeigen

30 **Fig. 1 A:** Die Struktur eines erfindungsgemäßen HCV-RNA-Konstrukts Ganz oben ist eine schematische Darstellung der Struktur des kompletten parentalen HCV-Genoms gegeben mit den Positionen der Gene für die Spaltungsprodukte core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A und NS5B innerhalb des Polyproteins, und den 5' und 3' nichttranslatierten Regionen (5' NTR und 3' NTR) — als Horizontalbalken dargestellt —, und mit den beiden für die Erzeugung der Subgenom-Konstrukte ausgewählten Positionen, nämlich der Position der 'GDD-katalytischen Domäne' der NS5B RNA Polymerase (GDD) und der Position der 3' Grenze der HCV-IRES (Nukleotidpositionen 1 bis 377 bzw. 1 bis 389) — oberhalb des Genomschemas eingezeichnet —. Die Zahlen unterhalb des Genomschemas bezeichnen die entsprechenden Nukleotidpositionen.

35 Darunter sind schematische Darstellungen der Strukturen zweier erfindungsgemäßer, modifizierter HCV-RNA-Konstrukte (Subgenom) gezeigt, bestehend aus der 5' HCV-IRES, dem Neomycinphosphotransferasegen (Neo^R), der EMCV-IRES (E-I) und den HCV Sequenzen von NS2 bzw. NS3 bis zum authentischen 3' Ende. Die Position der das NS5B Polymerase GDD-Motiv umfassenden 10 Aminosäuren-Deletion ist jeweils mit einem Dreieck (Δ) markiert.

45 **Fig. 1 B:** Das Ergebnis einer denaturierenden Formaldehyd-Agarose-Gelelektrophorese zum Nachweis von replizierter Plusstrang-RNA in transfizierten subpassagierten Huh-7 Zellklonen. Die Positionen der HCV-spezifischen RNAs (Pfeile) und der 28S rRNA sind rechts von Spur 12 angegeben, die Größen (Anzahlen der Nukleotide) der RNA-Marker (M) sind links von Spur 1 angegeben.

50 **Fig. 1 C** : Das Ergebnis eines PCR-Tests mit nachfolgendem Southern-Blot zum Nachweis der Abwesenheit von integrierter Replikon-DNA in den meisten der selektierten Zellklone. Spuren 1 und 2 zeigen die Positivkontrollen, Spur 13 die Negativ-Kontrolle. Die Zahlenangaben links der Spur 1 bezeichnen die Größe der Nukleotid-Marker-Moleküle.

55 **Fig. 2 A:** Das Ergebnis eines PCR-Tests mit nachfolgendem Southern-Blot zum sensitiven Ausschluß integrierter Replikon-DNA (Plasmid Moleküle I₃₇₇/NS3-3'/wt) in einem HCV-RNA-Konstrukt-haltigen Zellklon (9-13). Die Spuren 7 bis 11 repräsentieren das Ergebnis einer Titration von DNA-Molekülen des Konstrukts I₃₇₇/NS3-3'/wt ohne Zugabe von Gesamt-DNA des Zellklons 9-13 und die Spuren 2 - 6 repräsentieren die gleichen Plasmidmoleküle mit Zugabe von jeweils 1 µg 9-13 DNA vor der PCR (zwecks Ausschluß

eines Inhibitors der PCR in der DNA-Präparation). Spur 13 repräsentiert die Negativ-Kontrolle (PCR ohne DNA-Sonde). Spur 1 zeigt das Ergebnis, das mit einem µg Gesamt-DNA des Zellkons 9-13 erhalten wurde.

- 5 **Fig. 2 B:** Das Ergebnis eines Northern-Blot-Tests zur Quantifizierung von HCV Plus- und Minusstrang RNA. Die Pfeile markieren die Positionen von Replikon-RNA. Die "plus" und "minus" -Abgaben bezeichnen die positive (plus) bzw. negative (minus) Polarität der RNA-Kontrollen, die auf das Gel aufgetragen wurden. "Minusstrang" und "Plusstrang" bezeichnen die Spezifität der radioaktiven RNA-Sonden.
- 10 **Fig. 2 C:** Ergebnis einer Formaldehyd-Agarose-Gelelektrophorese nach radioaktiver Markierung der intrazellulär replizierten HCV-RNA zum Nachweis der Resistenz der HCV-RNA-Replikation gegen Dactinomycin.
- 15 **Fig. 3 A:** Nachweis von HCV-spezifischen Antigenen in den selektierten Zellklonen mittels Immunopräzipitation nach metabolischer Radioaktivmarkierung. Die Spuren 7 - 9 repräsentieren authentische Größenmarker (die nach transienter Expression eines HCV-RNA-Konstrukts in Huh-7-Zellen erhalten wurden); identifizierte HCV-Proteine sind am linken Rand von Spur 1 markiert, die Molekulargewichte (in Kilodalton) sind am rechten Rand von Spur 9 angegeben.
- 20 **Fig. 3 B:** Ergebnisse eines Immunfluoreszenztests zum Nachweis der subzellulären Lokalisation von HCV Antigenen.
- 25 **Fig. 4:** Schematische Darstellung der Struktur eines erfindungsgemäßen selektierbaren HCV-RNA-Konstrukts (komplettes Genom) bestehend aus der 5' HCV-IRES, dem Neomycinphosphotransferasegen (NeoR), einem heterologen IRES-Element, z.B. des Encephalomyocarditisvirus (E-I), dem vollständigen HCV-Leseraster und der authentischen 3' NTR.
- 30 **Fig. 5:** Schematische Darstellung der Struktur von HCV-RNA-Konstrukten mit insertiertem Antibiotikumresistenzgen (A) innerhalb der für das Polyprotein kodierenden Nukleotidsequenz (monocistronische RNA), und (B) innerhalb der 3' NTR (bicistronische RNA).
- 35 **Fig. 6:** Schematische Darstellung der Struktur von HCV-RNA-Konstrukten mit insertiertem Reportergen (A) als Teil eines HCV-Replikons von NS3 bis NS5B; — das Reportergen wird letztendlich durch virale oder durch zelluläre Proteasen aus dem Polyprotein gespalten und das selektierbare Markergen (Selektionsgen) bzw. das Resistenzgen durch Kontransfektion in die Zellen einschleust, (B) als Teil eines Fusionsgens aus Resistenz- und Reportergen (z.B. für die Neomycinphosphotransferase und green fluorescent Protein) (C) als Teil eines Replikons aus Resistenz- und Reportergen (z.B. für die Neomycinphosphotransferase und das green fluorescent Protein), die über eine Nukleotidsequenz verbunden sind, welche für eine Aminosäuresequenz kodiert (schraffierter Bereich), die von einer Protease gespalten werden kann oder die über eine selbstspaltende (autokatalytische) Aktivität verfügt, (D) als unabhängiges Gen (hier green fluorescent protein), das von einer eigenen internen Ribosomenbindungsstelle (IRES) aus exprimiert wird; — das Resistenzgen (hier: Neomycinphosphotransferase-Gen) wird davon unabhängig ebenfalls von einer eigenen internen Ribosomenbindungsstelle (IRES) aus exprimiert (polycistronisches Konstrukt).
- 40 **Fig. 7:** Schematische Darstellung der Struktur eines HCV-RNA-Konstrukts bei dem das Resistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA-Sequenz verbunden ist. Die dicken Linien stellen die HCV 5' und 3' NTRs dar, E-I ist eine heterologe interne Ribosomenbindungsstelle, die für die Expression des Resistenzgens notwendig ist, und das graue Quadrat stellt das Ribozym bzw. eine Erkennungsstelle für ein Ribozym dar.
- 45 **Fig. 8:** Schematische Darstellung der Struktur eines HCV-RNA-Konstrukts mit Resistenzgen und integriertem Fremdgen.
- 50 **Fig. 9:** Methodisches Vorgehen zum Vergleich der spezifischen Infektiosität (ausgedrückt als Anzahl gebildeter Zellkolonien) von Gesamt-RNA versus in vitro Transkripte. HCV-RNA wird mittels in vitro Transkription eines entsprechenden RNA-Konstrukts hergestellt und durch Messung der optischen Dichte bei 260 nm (OD 260 nm) quantifiziert. Eine definierte Anzahl dieser Moleküle wird mit einer bestimmten Menge Gesamt-RNA von naiven Huh-7 Zellen gemischt und diese Mischung mit Hilfe der Elektroporation in

naïve Huh-7 Zellen eingeschleust. Parallel dazu wird die Gesamt-RNA eines Zellklons, der mit der in Figur 1 beschriebenen Methode hergestellt wurde, mit einem im Stand der Technik bekannten Verfahren isoliert und die Menge der darin enthaltenen HCV-RNA mittels Northern-blot unter Verwendung einer HCV-spezifischen RNA-Sonde und anschließender Quantifizierung mittels Phosphorimager bestimmt. Eine definierte Menge dieser Gesamt-RNA wird analog den in vitro Transkripten in naïve Huh-7 Zellen transfiziert. Diese Zellen in beiden Ansätzen werden danach einer G418-Selektion unterzogen und die Anzahl der gebildeten Kolonien durch Auszählen nach fixieren und anfärben mit Coomassie-Brilliant-Blau bestimmt. Zur Bestimmung der Transfektionseffizienz wird jedem Transfektionsansatz 1 µg eines Plasmids zugesetzt, das die Expression der Luziferase erlaubt. Ein Aliquot der transfizierten Zellen wird nach 24 Stunden geerntet und die Luziferaseaktivität im jeweiligen Zellysat bestimmt. Die Anzahl der Kolonien wird jeweils auf die Luziferaseexpression normiert.

Fig. 10: Sequenzanalyse der 9-13 Klone. Gesamt-RNA des Zellklons 9-13, der durch Transfektion des HCV-RNA-Konstrukts I377/NS3-3' entstand, wurde mit einem im Stand der Technik bekannten Verfahren isoliert und das HCV-RNA-Konstrukt von Nukleotidposition 59 bis 9386 mit Hilfe der 'long-distance RT-PCR' unter Verwendung der primer S59 und A9413 amplifiziert. Die PCR-Fragmente wurden kloniert und 11 Klone (genannt 9-13 A - K) vollständig sequenziert, wobei sich die Klone D und I, E und G sowie H und J als identisch erwiesen. Die Positionen der Aminosäureunterschiede in der NS3-5B Region zwischen den reklonierten HCV-RNAs und dem parentalen Konstrukt sind mit einem dicken vertikalen Strich beim jeweiligen Klon markiert. Jeder Klon wurde mit dem Restriktionsenzym *Sfi* I verdaut und das jeweilige Fragment in das parentale Konstrukt inseriert. Diese Klone wurden jeweils in Huh-7 Zellen transfiziert und die Zellen wie in Figur 1 beschrieben einer Selektion unterzogen. Die Anzahl der mit jedem Konstrukt erhaltenen Zellklone ist rechts neben dem jeweiligen Konstrukt vermerkt.

Fig. 11 A: Prinzip der Replikationsbestimmung mit Hilfe eines Reportergens. Im oberen Teil der Figur ist das HCV-DNA-Konstrukt I₃₈₉/Luc/NS3-3' dargestellt, bestehend aus der HCV 5' NTR (Nukleotidposition 1-389), dem Luziferasegen (*luc*), der IRES des Encephalomyocarditis Virus, dem HCV NS3-5B und der 3' NTR. Die Position des aktiven Zentrums der NS5B RNA-Polymerase, in das ein inaktivierender Aminosäureaustausch eingeführt wurde, ist mit 'GND' angedeutet. Die Plasmide, die für das replikationskompetente bzw. das defekte HCV-RNA-Konstrukt kodieren, werden mit dem Restriktionsenzym *Sca* I verdaut und in eine in vitro Transkription mit der T7 RNA-Polymerase eingesetzt. Nach Entfernung der Matrizen-DNA werden die jeweiligen HCV-RNA-Konstrukte mittels Elektroporation in naive Huh-7 Zellen eingeschleust und diese in regelmäßigen Abständen geerntet.

Fig. 11 B: Vergleich der Luziferaseaktivitäten in Zellen transfiziert mit dem parentalen HCV-RNA-Konstrukt I₃₈₉/Luc/NS3-3'/wt (wt) oder den folgenden Varianten: Der inaktiven RNA (318 DN), der Variante 9-13F oder der Variante 5.1. Die Zellen wurden 6 (nicht gezeigt), 24, 48, 72, 96, 120, 144 und 168 Stunden nach der Transfektion geerntet und die Luziferaseaktivitäten luminometrisch bestimmt.

Fig. 12: Selektionierbare HCV-Vollängengenome (Konstrukte I₃₈₉/core-3'/5.1 und I₃₈₉/core-3'/9-13F).

(A) Schematische Darstellung des Vollängenkonstrukts. Der Bereich zwischen den beiden ange-deuteten Erkennungsstellen für das Restriktionsenzym *Sfi* I entspricht den Sequenzen der hoch-adaptierten RNA-Varianten 5.1. oder 9-13F.

(B) Anzahl der Kolonien die nach Transfektion von jeweils 0,1 µg in vitro transkribierter RNA der unter A dargestellten Konstrukte I₃₈₉/core-3'/5.1 in HUH7-Zellen erhalten wurden. Angegeben ist das Ergebnis eines repräsentativen Experimentes.

(C) Nachweis autonom replizierender HCV-Vollängen-RNAs in G418-resistenten Zellklonen, die nach Transfektion des entsprechenden in vitro Transkripts erhalten wurden. Die Abbildung zeigt das Autoradiogramm eines Northern Blots, der mit einer Sonde gegen das *neo*-Resistenzgen und der HCV 5' NTR hybridisiert wurde. Die in Spur 1 und 2 dargestellten Kontrollen entsprechen jeweils 10⁸ Molekülen der angegebenen in vitro Transkripte, gemischt mit Gesamt-RNA aus naiven Huh-7 Zellen. Die Negativkontrolle enthält ausschließlich Gesamt RNA aus naiven Huh-7 Zellen (Spur 3). Die Spuren 4-9 enthalten 3-10 µg Gesamt-RNA aus G418-resistenten Zellklonen, die nach Transfektion von in vitro transkribierter I₃₈₉/core-3'/5.1-RNA bzw. I₃₈₉/core-3'/9-13F-RNA erhalten wurden. Die für die Selektion verwendete G418-Konzentration ist jeweils angegeben. Fünf der dargestellten Zellklone enthalten die hoch adaptierte RNA-Variante 5.1 (Spur 4-8), einer die adaptierte RNA-Variante 9-13F (Spur 9).

Fig. 13: HCV-RNA-Konstrukte mit einem Reportergen. (A) Bicistronische HCV-RNA-Konstrukte. Das Reportergen wird mit Hilfe einer separaten IRES translatiert. (B) Monocistronische HCV-RNA-Konstrukte. Das Reportergenprodukt wird als Fusionsprotein mit einem HCV-Protein exprimiert. Die beiden Anteile sind über eine Erkennungssequenz für eine virale oder zelluläre Protease verbunden, die eine proteolytische Trennung der beiden fusionierten Proteinanteile erlaubt. Im gezeigten Beispiel wurden das Reportergenprodukt und das jeweilige HCV-Protein über eine Erkennungssequenz für Ubiquitin (Ub) fusioniert.

Fig. 14: Tricistronisches Vollängen HCV-RNA-Konstrukt, das zusätzlich zum Resistenzgen ein Fremdgen inseriert besitzt.

Fig. 15: Monocistronische HCV-RNA-Konstrukte, bei denen das Resistenzgenprodukt als Fusionsprotein mit dem HCV-Anteil exprimiert wird. Das Resistenzgen (RG) ist entweder als Fusionsprotein aktiv oder es wird so mit einer proteolytisch spaltbaren Sequenz mit dem HCV-Anteil fusioniert, daß das Resistenzgenprodukt durch eine zelluläre oder virale Protease vom HCV-Anteil abgespalten wird. Im gezeigten Beispiel wurde das Resistenzgen über die für Ubiquitin (Ub) kodierende Sequenz mit dem jeweiligen HCV-Anteil fusioniert.

Beispiel 1: Herstellung von HCV-RNA-Konstrukten

(A) Synthese und Klonierung eines vollständigen HCV-Konsensusgenoms mittels RT-PCR

[0053] Aus der Leber eines chronisch infizierten Patienten wurde das HCV-Genom, d.h. die HCV-RNA wie nachfolgend beschrieben isoliert:

[0054] Aus ca. 100 mg Leber wurde die komplette RNA gemäß dem Verfahren von Chomczynski und Sacchi (1987, Anal. Biochem. 162, 156) isoliert. Mit 1 µg dieser isolierten RNA wurde eine reverse Transkription mit den Primern A6103 (GCTATCAGCCGGTTCATCCACTGC) oder A9413 (CAGGATGGCCTATTGG CTGGAG) und dem 'expand reverse transcriptase'-System (Boehringer Mannheim, Deutschland) nach den Vorschriften des Herstellers durchgeführt. Mit den Produkten dieser reversen Transkription (RT) wurde eine Polymerase-Kettenreaktion (PCR=polymerase chain reaction) durchgeführt, und zwar unter Verwendung des 'expand long template'-Systems (Boehringer Mannheim, Deutschland), wobei der Puffer mit 2% Dimethylsulfoxid-Gehalt eingesetzt wurde. Nach einer Stunde bei 42°C wurde 1/8 dieses Reaktionsansatzes in einem ersten PCR-Durchgang mit den Primern A6103 und S59 (TGTCTTCACGCA-GAAAGCGTCTAG) oder A9413 und S4542 (GATGAGCT CGCCGCGAAGCTGTCC) eingesetzt. Nach 40 Zyklen wurde 1/10 dieses Reaktionsansatzes in einem zweiten PCR-Durchgang mit den Primern S59 und A4919 (AGCACAGCCCGCGTCATAGCACTCG) oder S4542 und A9386 (TTAGCTCCCCG TTCATCGGTTGG) eingesetzt. Nach 30 Zyklen wurden die PCR-Produkte mittels präparativer Agarose-Gel-Elektrophorese gereinigt und die dabei eluierten Fragmente wurden in den Vektor pCR2.1 (Invitrogen) oder pBSK II (Stratagene) ligiert. Vier Klone von jedem Fragment wurden analysiert und sequenziert, und es wurde eine Konsensus-Sequenz ermittelt. Zu diesem Zweck wurden die DNA-Sequenzen miteinander verglichen. Die Positionen, an denen sich die Sequenz eines der Fragmente von den übrigen unterschied, wurde als unerwünschte Mutation betrachtet. Im Fall von Mehrdeutigkeiten der Sequenz wurden kürzere sich überlappende PCR-Fragmente der betreffenden Region amplifiziert und mehrere Klone sequenziert. Auf diese Weise konnten zahlreiche potentielle Mutationen in jedem Fragment identifiziert und somit eine isolat-spezifische Konsensussequenz etabliert werden. Diese etablierte Konsensussequenz bzw. dieses Genom gehört zum weltweit verbreiteten Genotyp 1b. Die nicht translatierte Region am 3'-Ende (=3' NTR) wurde mittels konventioneller PCR erhalten, wobei ein Antisense-Primer eingesetzt wurde, der die letzten 24 Nukleotide des im Stand der Technik bekannten 'X-tails' (Tanaka et al., 1995, Biochem. Biophys. Res. Commun. 215, 744; und Rice, PCT/US 96/14033) abdeckt. Die authentische nicht translatierte Region am 5'-Ende (=5' NTR) strangabwärts vom T7 Promotor wurde mittels PCR erzeugt, wobei zum einen ein Oligonukleotid verwendet wurde, das einem verkürzten T7 Promotor (TAA TAC GAC TCA CTA TAG) und den ersten 88 Nukleotiden von HCV entspricht, und zum anderen eines der vorgenannten Plasmide eingesetzt wurde, das eines der 5' Fragmente des Genoms trägt. Aus den subgenomischen Fragmenten mit der geringsten Anzahl an Nicht-Konsensus-Austauschen wurde ein komplettes HCV-Konsensusgenom zusammengesetzt und in einen modifizierten pBR322-Vektor inseriert. Abweichungen von der Konsensussequenz wurden mittels ortsgerechter Mutagenese ("site-directed mutagenesis") beseitigt. Um "run-off"-Transkripte mit einem authentischen 3' Ende herzustellen, wurde die 3'-NTR der Isolate (mit dem Ende TGT) zu AGT modifiziert (gemäß der Sequenz vom Genotyp 3 = Klon 'WS' nach Kolykhalov et al., 1996, J. Virol. 70, 3363) und außerdem wurde ein zusätzlicher Nukleotidaustausch an Position 9562 vorgenommen, um die A:T Basenpaarung in der Haarnadelstruktur am 3' Ende der 3' NTR (Kolykhalov et al. ibid.) beizubehalten. Um eine interne Restriktionsstelle für das Enzym Scal zu beseitigen, wurde ferner ein sog. stiller ("silent") Nukleotidaustausch vorgenommen. Nach dem Zusammenfügen des Vollängen-Genoms mit passenden 5'- und 3' NTRen wurde die komplette HCV-Sequenz überprüft. Dabei wurde kein ungewünschter Nukleotidaustausch

gefunden.

[0055] Das auf diese Weise hergestellte HCV-Genom sollte per Definition hepatotrop sein.

(B) Synthese selektierbarer HCV-Subgenom-Konstrukte

5

[0056] Unter Verwendung des unter (A) beschriebenen Konsensusgenoms wurden HCV-Subgenom-Konstrukte hergestellt, die das Antibiotikumresistenzgen Neomycin-Phosphotransferase (NPT) und zwei Sequenzen von internen Ribosomenbindungsstellen (IRES) enthalten. Die hierfür angewendeten biochemischen Verfahrenstechniken sind dem Fachmann bekannt und geläufig (siehe: Sambrook, J., E.F. Fritsch, T. Maniatis, 1989, *Molecularcloning: a laboratory manual*, 2nd ed., Cold Spring Harbour Laboratory, Cold Spring Harbor, N.Y.; Ausubel et al. (eds.), 1994, *Current Protocols in Molecular Biology*, Vol. 1-3, John Wiley & Sons Inc., New York). Das Antibiotikumresistenzgen wurde unmittelbar hinter der 5' NTR inseriert, wodurch eine bicistronische RNA erhalten wurde (siehe Fig. 1 A). Ebenso gut kann das Antibiotikumresistenzgen aber auch an anderer Stelle des HCV-Subgenom-Konstrukts inseriert werden, beispielsweise innerhalb der für das Polyprotein kodierenden Nukleotidsequenz, wodurch eine monocistronische RNA erhalten wird (siehe Fig. 5 A) oder in die 3' NTR (siehe Fig. 5 B). Bei den IRES-Elementen handelt es sich zum einen um eine der beiden HCV-IRES-Varianten Nukleotide 1-377 oder Nukleotide 1-389, und zum anderen um die IRES des Enzephalomyocarditis Virus, die die Translation der HCV Sequenz strangabwärts von den Genen für NS2 oder NS3 bis zu dem authentischen 3' Ende des Genoms steuert.

[0057] Die beiden genannten HCV-IRES-Varianten wurden wie folgt ermittelt: Auf der Basis von Deletionsanalysen der 3' Grenze der HCV-IRES (Reynolds et al. 1995, *EMBO J.* 14, 6010) wurden verschiedene Abschnitte der 5' NTR mit dem NPT Gen fusioniert und anhand von Kofusionen mit einem das T7 RNA Polymerase Gen enthaltenden Plasmid hinsichtlich der maximalen Anzahl gebildeter Kolonien analysiert. Die besten Ergebnisse wurden mit den HCV Sequenzen von 1-377 und 1-389 erhalten. Da sich das AUG-Startkodon des HCV Polyproteins an Position 342 befindet und somit in der IRES-Sequenz enthalten ist, kommt es zu einer Fusion von 12 bzw. 16 Aminosäuren des HCV-Kapsidproteins ("Core-Proteins") mit der Neomycin Phosphotransferase (siehe Fig. 1 A).

[0058] Diese modifizierten HCV-Subgenom-Konstrukte erhielten dementsprechend die Bezeichnungen I₃₇₇/NS2-3' (oder I₃₇₇/NS3-3') und I₃₈₉/NS2-3' (oder I₃₈₉/NS3-3'). Sie sind in Fig. 1A schematisch dargestellt.

[0059] Mit in-vitro-Transkripten dieser modifizierten parentalen HCV-Subgenom-Konstrukte I₃₇₇/NS2-3' (oder I₃₇₇/NS3-3') und I₃₈₉/NS2-3' (oder I₃₈₉/NS3-3') wurden verschiedene Zelllinien und Primärzellkulturen von menschlichen Hepatocyten transfiziert.

[0060] Als parallele Negativ-Kontrolle zu allen Transfektionsexperimenten wurde zu jedem modifizierten parentalen HCV-Subgenom-Konstrukt ein entsprechend modifiziertes aber defektes Subgenom konstruiert, das sich von dem parentalen dadurch unterscheidet, daß es innerhalb des Leserasters eine Deletion von 10 Aminosäuren aufweist, die das aktive Zentrum der NS5B RNA Polymerase umfaßt (Behrens et al., 1996, *EMBO J.* 15, 12; und Lohmann et al., 1997, *J. Virol.* 71, 8416).

(C) Synthese selektierbarer HCV-Genom-Konstrukte

[0061] Ein NS2-3' Subgenomkonstrukt, das am 5' Ende mit einem Fragment des Luziferasegens und der vollständigen EMCV-IRES verbunden ist, wurde mit NcoI und SpeI restringiert und mittels präparativer Agarosegelelektrophorese gereinigt. Der so erhaltene Vektor wurde in einer 3-Faktor Ligation mit einem NcoI/NotI-HCV-Fragment, entsprechend den Nukleotidpositionen 342 bis 1968 des HCV-Genoms und mit einem NotI/SpeI-Fragment, entsprechend den Nukleotidpositionen 1968-9605 ligiert. Das entstandene Konstrukt, bei dem das vollständige HCV-Leseraster und die 3' NTR stromabwärts dem Luziferasegenfragment und der EMCV-IRES liegen, wurde danach mit PmeI und SpeI restringiert und mit dem analog restringierten I₃₈₉/NS3-3'/wt-Subgenomkonstrukt-Vektor ligiert. Dieses selektionierbare HCV-Genomkonstrukt ist in Fig. 4 dargestellt.

(D) Herstellung von den HCV-RNA-Konstrukten entsprechenden in-vitro-Transkripten

[0062] Die vorstehend beschriebenen gereinigten Plasmid DNAs wurden mit Scal linearisiert und nach Phenol/Chloroform-Extraktion und Isopropanol-Präzipitation in eine In-vitro-Transkriptionsreaktion eingesetzt unter Verwendung der folgenden Komponenten: 80 mM HEPES, pH 7.5, 12,5 mM MgCl₂, 2 mM Spermidin, 40 mM Dithiothreitol, 2 mM von jedem NTP, 1 Einheit RNasin/µl, 50 µg/ml restringierte DNA und ca. 2 Einheiten/µl T7 RNA Polymerase. Nach 2 Std. bei 37°C wurde die Hälfte der Menge an T7 Polymerase zugegeben und der Reaktionsansatz weitere 2h inkubiert. Zur Entfernung von DNA wurde die Mischung mit saurem Phenol extrahiert (U. Kedzierski, J.C. Porte, 1991, *Bio Techniques* 10, 210), mit Isopropanol präzipitiert, das Pellet in Wasser gelöst und mit DNase (2 Einheiten pro µg DNA) für 60 Min. bei 37°C inkubiert. Nach anschließender Extraktion mit saurem Phenol, saurem Phenol/Chloroform und Chloroform und Isopropanol-Präzipitation wurde die gelöste RNA mittels optischer Dichtemessungen quantifiziert und

ihre Unversehrtheit mittels Formaldehyd-Agarose-Gelelektrophorese überprüft.

Beispiel 2: Transfektionsexperimente mit der Hepatomazelllinie Huh-7

[0063] Bei sämtlichen Transfektionsexperimenten wurde sorgfältig darauf geachtet, daß jegliche Matrizen-DNA zuvor entfernt worden war, um zu vermeiden, daß solche DNA in transfizierte Zellen integrieren und diesen unabhängig von einer HCV-Replikation eine Neomycin-Resistenz vermitteln konnte. Deshalb wurde im Anschluß an die in-vitro-Transkription (Beispiel 1 D) die Reaktionsmischung mit 2 Einheiten DNase pro µg DNA für 60 Min. bei 37°C behandelt und mit saurem Phenol, saurem Phenol/Chloroform und Chloroform extrahiert. Vor der Verwendung für die Transfektion wurde die präzipitierte RNA mittels Formaldehyd Agarose Gel Elektrophorese analysiert.

[0064] Es wurden drei separate Transfektionsexperimente mit der hoch differenzierten humanen Hepatomazelllinie Huh-7 (gemäß Nakabayashi et al. 1982, *Cancer Res.* 42, 3858) durchgeführt. Dabei wurde jeweils 15 µg RNA in 8×10^6 Huh-7-Zellen mit Hilfe der Elektroporation eingebracht und diese Zellen anschließend in Kulturschalen von 10 cm Durchmesser ausgesät. 24 Stunden nach der Aussaat wurde Neomycin (= G418) in einer Endkonzentration von 1 mg/ml zugegeben. Das Kulturmedium wurde zweimal pro Woche gewechselt. Nach 3 - 5 Wochen waren kleine Kolonien erkennbar, die isoliert und unter den gleichen Kulturbedingungen passagiert wurden.

[0065] Die Zellklone, die im Verlauf des ersten Experiments erhalten wurden, wurden isoliert und subpassagiert. Während dieser Prozedur starben die meisten Klone und die Endausbeute betrug nur noch 9 Klone von Zellen, die mit den parentalen HCV-Subgenom-Konstrukten transfiziert worden waren und 1 Klon (Klon 8-1) von Zellen, die mit einem defekten HCV-Genom-Konstrukt, nämlich einer defekten NS2-3' HCV-RNA transfiziert worden waren. Außer einer verkürzten Verdopplungszeit und dem gelegentlichen Auftreten von irregulär geformten Zellen wurden keine beständigen morphologischen Unterschiede zwischen diesen 9 Zellklonen und dem einen Zellklon (Klon 8-1) oder den parentalen Huh-7 Zellen gefunden.

[0066] Die Hauptkriterien für funktionierende HCV-Genomkonstrukte sind die Bildung von viraler RNA mit korrekter Größe und die Abwesenheit von (integrierter) Plasmid DNA, die eine G418-Resistenz übertragen bzw. vermitteln könnte.

[0067] Um die HCV-RNA in den Huh-7-Zellen zu bestimmen, wurde die Gesamt-RNA isoliert und mittels des gängigen Northern-Blot Verfahrens unter Verwendung einer Plusstrang-spezifischen Ribosonde (= RNA-Sonde) analysiert. Hierfür wurde von den jeweiligen Zellklonen Gesamt-RNA nach der Methode von Chomczynski und Sacchi 1987, *Anal. Biochem.* 162, 156 isoliert, und 10 µg RNA, was dem Gesamt-RNA-Gehalt von $0,5 - 1 \times 10^6$ Zellen entspricht, mittels denaturierender Formaldehyd-Agarose-Gelelektrophorese aufgetrennt (Spuren 3 bis 12 der Fig. 1 B). Als Größenmarker mit authentischer Sequenz wurden gleichzeitig 10^9 in-vitro-Transkripte (ivtr.), die zu den $I_{389}/NS2-3'/wt$ oder den $I_{389}/NS3-3'/wt$ Replikon-RNAs korrespondieren, mit aufgetrennt (Spur 1 bzw. Spur 2). Die aufgetrennte RNA wurde auf Nylon-Membranen transferiert und mit radioaktiv markierter Plusstrang-spezifischer RNA-Sonde, die komplementär zu dem kompletten NPT-Gen und der HCV-IRES von Nukleotid 377 bis Nukleotid 1 war, hybridisiert. Die Positionen der HCV-spezifischen RNAs (Pfeile) und der 28S rRNA sind rechts von Spur 12 angegeben, die Größen (Anzahlen der Nukleotide) der RNA-Marker sind links von Spur 1 angegeben. Die RNA Marker-Fragmente enthalten HCV-Sequenzen und hybridisieren deshalb mit der Ribosonde (= RNA-Sonde). Die Ergebnisse dieser Analyse sind in Fig. 1 B dargestellt.

[0068] Mit Ausnahme des mit dem defekten HCV-Genom-Konstrukt transfizierten Klons 8-1, lieferten alle Zellklone homogene HCV-RNAs korrekter Länge (ca. 8640 Nukleotide im Fall des NS2-3' und ca. 7970 Nukleotide im Fall des NS3-3' Replikons). Dieser Befund ist ein Indiz dafür, daß die funktionalen Replikons bzw. die funktionalen HCV-Genom-Konstrukte die G418 Resistenz übertragen. Um auszuschließen, daß die G418 Resistenz auf eine Plasmid-DNA zurückzuführen ist, die in das Genom der Huh-7 Wirtszelle integriert ist und unter der Kontrolle eines zellulären Promotors transkribiert wird, wurde von jedem Klon die DNA mittels einer NPT-Gen-spezifischen PCR untersucht. Hierbei wurde aus den selektierten Huh-7-Zellklonen die DNA mittels Verdau mit Proteinase K (40 µg/ml, 1h, 37°C) in 10mM Tris, pH7,5, 1mM EDTA, 0,5% SDS und anschließender Extraktion mit Phenol, Phenol/Chloroform und Isopropanolpräzipitation isoliert. Das DNA-Präzipitat wurde in 10 mM Tris (pH 7,5) und 1 mM EDTA gelöst und 1 Stunde mit Rnase A inkubiert. Im Anschluß an eine Phenol/Chloroform Extraktion und Ethanol Präzipitation wurde 1 µg DNA, entsprechend $4 - 8 \times 10^4$ Zellen, mittels PCR unter Einsatz NPT-Gen-spezifischer Primer (5'-TCAAGACCGACCTG TCCGGTGCCC-3' und 5'-CTTGAGCCTGGCGAACAGTTCGGC-3') analysiert und ein DNA-Fragment bestehend aus 379 Nukleotiden erzeugt. Die Spezifität des PCR-Produkts wurde mittels Southern Blot Verfahren nachgewiesen, wobei ein Digoxigenin-markiertes DNA Fragment eingesetzt wurde, das zu dem NPT-Gen korrespondiert. Als Positiv-Kontrollen (zum Nachweis etwa vorhandener kontaminierender Nukleinsäuren) wurde das PCR-Verfahren mit 10^7 Plasmid Molekülen oder 1 µg DNA aus einer BHK Zelllinie, die stabil mit einem Neomycin-Resistenz-Gen transfiziert war, durchgeführt, und als Negativ-Kontrolle wurde die PCR mit denselben Reagenzien aber ohne zugesetzte DNA durchgeführt.

Die Ergebnisse dieser Untersuchung sind in Fig. 1 C dargestellt. Die Spuren 1 und 2 repräsentieren die Positiv-Kontrol-

len, Spur 13 repräsentiert die Negativ-Kontrolle. Die Zahlenangaben links der Spur 1 bezeichnen die Größe der Nukleotid-Marker-Moleküle. Außer in Klon 7-3 (Fig. 1C, Spur 3), der von Zellen nach Transfektion mit einem NS2-3' Replikon/NS2-3'HCV-Genom-Konstrukt stammt, und in Klon 8-1 (Fig. 1C, Spur 12), der von Zellen nach Transfektion mit einem defekten HCV-Genom-Konstrukt stammt, war in keinem Zellklon eine NPT-DNA nachweisbar. Dieser Befund ist ein weiteres Indiz dafür, daß die G418 Resistenz der meisten Klone durch die replizierende HCV-RNA vermittelt wurde. Aber auch unabhängig von diesen Ergebnissen ist es unwahrscheinlich, daß HCV-RNAs mit korrekter Größe von integrierter Plasmid DNA erzeugt wird, denn die für die in-vitro-Transkription verwendeten Plasmide enthalten weder einen eukaryontischen Promotor noch ein Polyadenylierungssignal. Im Fall des Klons 7-3 ist die Resistenz deshalb höchst wahrscheinlich sowohl durch das HCV-RNA-Konstrukt bzw. die replizierende HCV-RNA als auch durch eine integrierte NPT DNA Sequenz vermittelt worden, während die Resistenz der Zellen von Klon 8-1 ausschließlich auf die integrierte Plasmid DNA zurückzuführen ist.

[0069] Um zu bestätigen, daß die G418 Resistenz von einer autonom replizierenden HCV-RNA vermittelt ist, wurde der Klon 9-13 (Fig. 1 B, Spur 11) weiteren Tests unterworfen. Klon 8-1, der integrierte Kopien des NPT-Gens trägt, wurde überall als Negativkontrolle eingesetzt. Mit dem Ziel, die Anwesenheit von NPT-DNA im Klon 9-13 rigoros auszuschließen, wurde eine PCR durchgeführt, die den Nachweis von < 1000 NPT-Gen-Kopien in ~ 40.000 Zellen erlaubt. Das Ergebnis dieser PCR ist in Fig. 2A dargestellt. Im einzelnen wurde bei dieser PCR wie folgt verfahren:

Es wurden jeweils 10^6 - 10^2 Plasmid Moleküle ($I_{377}/NS3-3'/wt$) entweder direkt (Spuren 7 - 11) oder nach Zugabe von jeweils $1 \mu g$ 9-13 DNA (Spuren 2 - 6) in dem Test eingesetzt. Die Spezifität der amplifizierten DNA Fragment wurde mittels Southern Blot unter Verwendung einer NPT-spezifischen Sonde bestimmt. Eine PCR ohne DNA-Sonde wurde als Negativ-Kontrolle durchgeführt (Spur 12).

Selbst mit dieser sensitiven Methode wurde in einem μg DNA des Zellklons 9-13 keine Plasmid DNA gefunden (Spur 1). Um die Menge an HCV Plus- und Minusstrang RNAs in diesen Zellen abzuschätzen, wurde eine Verdünnungsreihe von Gesamt-RNA mit dem Northern-Blot-Verfahren unter Verwendung einer Plus- oder Minusstrang-spezifischen radioaktiv markierten Ribosonde (= RNA-Sonde) analysiert. Hierfür wurden jeweils 8, 4 oder $2 \mu g$ Gesamt-RNA, die aus den Zellklonen 9-13 und 8-1 isoliert worden waren, parallel zu bekannten Mengen analoger in-vitro-Transkripte mit Plus- oder Minusstrang-Polarität (Kontroll-RNAs) im Northern-Blot-Verfahren analysiert und anschließend einer Hybridisierung unterworfen. Die Hybridisierung wurde mit einer Plusstrang-spezifischen Ribosonde, die das komplette NPT-Gen und die HCV-IRES abdeckte ('plusstrang', obere Bildtafel), oder mit einer Minusstrang-spezifischen RNA-Sonde, die zu der NS3-Sequenz komplementär war ('minusstrang', untere Bildtafel) durchgeführt. Die Pfeile markieren die Positionen von Replikon-RNA. Die Ergebnisse dieser Analyse sind in Fig. 2 B dargestellt.

Im Fall des Plusstrangs wurden ca. 10^8 Kopien/ μg Gesamt-RNA nachgewiesen, was 1000 - 5000 HCV-RNA-Molekülen pro Zelle entspricht, während die Menge an Minusstrang-RNA 5- bis 10-fach niedriger war. Dieses Ergebnis stimmt mit der Annahme überein, daß die Minusstrang RNA die replikative Zwischenform bzw. Zwischenkopie ist, die als Vorlage für die Synthese der Plusstrang Moleküle dient.

Da die Reaktion im wesentlichen von der viralen RNA-abhängigen RNA Polymerase katalysiert wird, sollte die Synthese der HCV-RNAs resistent gegen Dactinomycin sein, einem Antibiotikum, das selektiv die RNA-Synthese von DNA-Matrizen inhibiert, nicht jedoch die RNA-Synthese von RNA-Matrizen. Um diese Vermutung zu bestätigen, wurden Zellen mit [3H] Uridin in Anwesenheit von Dactinomycin inkubiert, die radioaktiv markierten RNAs extrahiert, mittels denaturierender Agarose-Gel-Elektrophorese aufgetrennt und mit Hilfe eines handelsüblichen Bio-Imagers unter Verwendung einer [3H]-sensitiven Bildplatte analysiert. Hierfür wurden jeweils ca. 5×10^5 Zellen der Klone 9-13 und 8-1 mit $100 \mu Ci$ [3H]Uridin für 16 Std. in Abwesenheit (-) oder Gegenwart (+) von $4 \mu g/ml$ Dactinomycin (Dact) inkubiert. Im Anschluß an diese Markierungsreaktion wurde die Gesamt-RNA präpariert und mittels Formaldehyd-Agarose-Gel-Elektrophorese analysiert. In den beiden ersten Spuren ist nur 1/10 der Gesamt-RNA dargestellt. Die radioaktiv markierte RNA wurden mit einem BAS-2500 Bio-Imager (Firma Fuji) sichtbar gemacht.

Die Ergebnisse dieser Analyse sind in Fig. 2 C dargestellt. In Übereinstimmung mit dem Inhibitor-Profil der NS5B Polymerase (Behrens et al., 1996, *EMBOJ.* 15, 12 und Lohmann et al., 1997, *J Virol.* 71, 8416) war die Replikation der HCV RNA nicht durch Dactinomycin beeinflusst worden, während die Synthese von zellulärer RNA gehemmt worden war. Um die Identität der viralen RNA zu bestätigen, wurde eine RT-PCR zur Reklonierung der replizierten Sequenzen durchgeführt. Die Sequenzanalyse der reklonierten RNA zeigte, daß die RNA in dem Klon 9-13 HCV-spezifisch ist und mit dem transkribierten Transkript des HCV-Konstrukts $I_{377}/NS3-3'/wt$ übereinstimmt.

[0070] Zur Analyse der viralen Proteine wurden die betreffenden Zellen zunächst metabolisch mit [^{35}S] Methionin/Cystein radioaktiv markiert, anschließend lysiert und danach die HCV-spezifischen Proteine mittels Immunopräzipitation aus den Zell-Lysaten isoliert. Die Ergebnisse dieser Analysen sind in Fig. 3 A dargestellt. Im einzelnen wurde dabei wie folgt verfahren: Zellen der Zellklone 9-13 (wt) und 8-1 (Δ) waren durch Behandlung für 16 Stunden mit einer dem Fachmann geläufigen und im Handel erhältlichen Protein-Markierungs-Mischung (z.B. NEN Life Science) metabolisch radioaktiv markiert worden. Mittels Immunopräzipitation (IP) unter nicht-denaturierenden Bedingungen (z.B. nach Bartenschlager et al., 1995, *J. Virol.* 69, 7519) und unter Verwendung von drei verschiedenen Antiseren (3/4, 5A, 5B, gemäß Markierung am oberen Ende der Spuren 1 bis 12) waren die HCV-spezifischen Proteine vom Zell-Lysat abge-

trennt worden.. Die Immunokomplexe wurden mittels Tricine SDS-PAGE analysiert und mittels Autoradiographie sichtbar gemacht. Um authentische Größenmarker zu erhalten, wurde das homologe Replikonkonstrukt I₃₇₇/NS3-3'/wt einer transienten Expression mit dem Vaccinia Virus T7-Hybrid System in Huh-7 Zellen unterworfen. Die dabei erhaltenen Produkte waren als Größenmarker (Spuren 7 - 9) parallel zu den Zellen der Klone 9-13 und 8-1 behandelt worden. Identifizierte HCV-Proteine sind am linken Rand von Spur 1 markiert, die Molekulargewichte (in Kilodalton) sind am rechten Rand von Spur 9 angegeben. Es ist anzumerken, daß das verwendete NS3/4-spezifische Antiserum ('3/4') bevorzugt mit NS4A und NS4B reagiert, was zu einer Unterrepräsentation von NS3 führt.

[0071] Alle viralen Antigene waren eindeutig nachweisbar und ihre apparenten Molekulargewichte zeigten keine Abweichungen gegenüber denjenigen, die nach transients Expression desselben bicistronischen HCV-RNA-Konstrukts in den ursprünglichen Huh-7 Zellen ermittelt wurden. Um die subzelluläre Verteilung der viralen Antigene zu bestimmen, wurde eine Immunofluoreszenz-Nachweisreaktion unter Einsatz von NS3- und NS5A-spezifischen Antisera durchgeführt (z.B. nach Bartenschlager et al., 1995, *J. Virol.* 69, 7519). Hierfür wurden Zellen der Klone 9-13 (wt) und 8-1 (Δ) 24 Std. nach dem Aussäen auf Deckgläsern mit Methanol/Azeton fixiert und mit polyklonalen NS3- oder NS5A-spezifischen Antisera inkubiert. Die gebundenen Antikörper wurden mit einem kommerziell erhältlichen FITC-konjugierten Anti-Kaninchen-Antiserum sichtbar gemacht. Zur Unterdrückung unspezifischer Fluoreszenzsignale wurden die Zellen mit dem Farbstoff 'Evans Blue' gegengefärbt.

[0072] Die Ergebnisse dieses Nachweistests sind in Fig. 3 B dargestellt. Mit beiden Antisera war eine starke Fluoreszenz im Zytoplasma nachweisbar. Die NS5A-spezifischen Antisera führten außerdem zu einer schwachen Zellkern-Fluoreszenz, was darauf hindeutet, daß zumindest kleine Mengen dieses Antigens auch zum Zellkern gelangen. Die generell dominierende Präsenz der viralen Antigene im Zytoplasma ist jedoch ein starkes Indiz dafür, daß die HCV-RNA Replikation im Zytoplasma stattfindet — so wie das bei den meisten RNA-Viren der Fall ist.

[0073] Diese Ergebnisse belegen klar, daß mit dem hier beschriebenen Versuchsansatz der Aufbau eines Zellsystems für das HCV gelungen ist, dessen Effizienz alles bisher bekannte um Größenordnungen übersteigt und erstmalig den Nachweis viraler Nukleinsäuren und Proteine mit konventionellen und bewährten biochemischen Methoden erlaubt. Erst diese Effizienz erlaubt überhaupt detaillierte Untersuchungen der HCV-Pathogenese, genetische Analysen verschiedener HCV-Funktionen und ein genaues Studium der Virus-/Wirtszellwechselwirkungen, wodurch sich neue Ansatzpunkte für die Entwicklung einer antiviralen Therapie definieren lassen.

Beispiel 3: Transfektion von Huh-7 Zellen mit HCV-Genomkonstrukten

[0074] Huh-7 Zellen werden wie in Beispiel 2 beschrieben transfiziert und selektioniert, wobei hier jedoch selektionierbare Konstrukte verwendet werden, die das vollständige Virusgenom enthalten. Die erhaltenen Zellklone werden analog dem Beispiel 2 mittels PCR auf Abwesenheit von HCV-DNA untersucht und die produktive Replikation der HCV-RNA wird danach mittels Northern Blot, [³H]Uridinmarkierung in Anwesenheit von Dactinomycin, Nachweis der viralen Proteine bzw. Antigene vorzugsweise mit Hilfe des Western Blots, der Immunopräzipitation oder der Immunfluoreszenz nachgewiesen. Im Gegensatz zu den im Beispiel 2 beschriebenen Ansätzen lassen sich mit dem hier beschriebenen Konstrukt außerdem vollständige und sehr wahrscheinlich infektiöse Viren erhalten, was bei den dort (in Beispiel 2) beschriebenen Subgenomkonstrukten nicht der Fall ist. Diese Viren, die in der Zelle und dem Zellkulturüberstand vorhanden sind, werden beispielsweise mittels Ultrazentrifugation, Immunpräzipitation oder Fällung mit Polyethylenglykol konzentriert und alle exogenen, d.h. nicht im Viruspartikel eingebauten Nukleinsäuren werden mittels Inkubation mit Nukleasen (RNase, DNase, Mikrococusunuklease) verdaut. Auf diese Weise lassen sich alle kontaminierenden Nukleinsäuren, die nicht im schützenden Viruspartikel enthalten sind, entfernen. Die geschützte virale RNA wird nach Inaktivierung der Nukleasen, beispielsweise mittels Inkubation mit Proteinase K in einem SDS-haltigen Puffer durch Extraktion mit Phenol und Phenol/Chloroform isoliert und mittels Northern Blot oder RT-PCR unter Verwendung HCV-spezifischer Primer nachgewiesen. Auch in diesem Versuchsansatz ist die Kombination des beschriebenen HCV-Konsensusgenoms mit einem Selektionsmarker entscheidend für die effiziente Produktion von viraler RNA, viralem Protein und damit von HCV-Partikeln.

Beispiel 4: Herstellung und Anwendung eines HCV-RNA Konstrukts, bei dem das Resistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-Subgenom-Sequenz verbunden ist.

[0075] Es wird ein HCV-RNA-Konstrukt gemäß Beispiel 1 oder Beispiel 3 hergestellt, bei dem ein Antibiotikumresistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA-Sequenz verbunden ist. Solche Konstrukte sind in Fig. 7 schematisch dargestellt. Huh-7 Zellen werden wie in Beispiel 2 beschrieben mit diesem HCV-RNA-Konstrukt transfiziert. Nach der Transfektion in die Zellen erfolgt zunächst die Selektion mit dem entsprechenden Antibiotikum. In den dabei erhaltenen Zellklonen wird das inklonierte Ribozym aktiviert oder, im Fall eines Konstrukts, das eine Erkennungsstelle für ein Ribozym trägt, wird das Ribozym in die Zelle eingeschleust (z.B. mittels Transfektion eines Ribozymkonstrukts oder Infektion mit einem viralen Expressionsvektor, in den das entsprechende

Ribozym eingesetzt wurde). In beiden Fällen wird durch die ribozymvermittelte Spaltung das Resistenzgen von der HCV-RNA-Sequenz abgetrennt. Das Ergebnis ist im Fall des HCV-Genom-Konstrukts ein authentisches HCV-Genom ohne Resistenzgen, das zur Bildung authentischer infektiöser Viruspartikel befähigt ist. Im Fall des HCV-Subgenom-Konstrukts entsteht ein HCV-Replikon ohne Resistenzgen.

Beispiel 5: Kotransfektion eines HCV-RNA-Konstrukts mit einem separaten Luziferase-Transfektionskonstrukt

[0076] Es wird ein HCV-RNA-Konstrukt gemäß Beispiel 1 (A) oder Beispiel 3 oder Beispiel 4 hergestellt. Parallel dazu wird ein Transfektionskonstrukt hergestellt, welches das Luziferasegen umfaßt, wobei dieses Luziferasegen ver-
 mittels einer ersten Nukleotidsequenz, die für eine HCV-Protease- (z.B. NS3-Protease-) Spaltungsstelle kodiert, mit
 einer zweiten Nukleotidsequenz, die für ein anderes Protein oder einen Teil eines anderen Proteins kodiert, verbunden
 ist. HCV-RNA-Konstrukt und Transfektionskonstrukt werden in beliebige Wirtszellen, vorzugsweise Hepatomazellen,
 insbesondere Huh-7-Zellen, eingeschleust. Dies kann auf die in Beispiel 2 beschriebene Art und Weise geschehen.
 Das Produkt des modifizierten Luziferasegens ist ein Luziferase-Fusionsprotein, in dem die Luziferase auf Grund der
 Fusion mit dem Fremdanteil inaktiv ist. In transfizierten Zellen mit hoher HCV-Replikation wird das Fusionsprotein, das
 ja eine Schnittstelle für eine HCV-Protease enthält, gespalten und damit die aktive Form der Luziferase freigesetzt, die
 sich durch luminometrische Messung bestimmen läßt. Wird die Replikation des HCV-RNA-Konstrukts gehemmt, wird
 das Fusionsprotein nicht gespalten und keine aktive Luziferase freigesetzt. Infolgedessen ist die quantitative Bestim-
 mung der Luziferase ein Maß für die Replikation des HCV-Subgenom-Konstrukts. Anstelle des Luziferasegens kann
 ebenso ein anderes Reportergen verwendet werden, das in analoger Weise modifiziert ist, so daß seine Expression
 von der Virusreplikation abhängt, obwohl dieses Reportergen nicht Bestandteil des HCV-Subgenom-Konstrukts ist. Es
 kann auch ein zelluläres Protein, welches durch die HCV-Proteine oder Nukleinsäure inaktiviert oder aktiviert wird, als
 sogenannter Surrogatmarker verwendet werden. In diesem Fall ist die Expression bzw. Aktivität dieses Surrogatmark-
 ers ein Maß für die Replikation der viralen DNA.

Beispiel 6: Herstellung von HCV-Subgenom-Konstrukten mit integrierten Fremdgenen zur Verwendung als leberzellspezifische Genfährten für die Gentherapie

[0077] Diese rekombinanten und selektionierbaren HCV-Subgenom-Konstrukte werden in trans-komplementie-
 rende Helferzelllinien transfiziert, d.h. in Zelllinien, die induzierbar oder konstitutiv die fehlenden Funktionen (beispiels-
 weise die Strukturproteine) exprimieren. Zellklone, die ein funktionelles HCV-Subgenom-Konstrukt enthalten, lassen
 sich durch entsprechende Selektion etablieren. Die von der Wirtszelle exprimierten Virus-Strukturproteine erlauben die
 Bildung von Viruspartikeln, in die die RNA der HCV-Subgenom-Konstrukte eingeschleust wird. Das Ergebnis sind also
 virus-ähnliche Partikel, die ein erfindungsgemäßes HCV-Subgenom-Konstrukt einschließlich des inklonierten Fremd-
 gens enthalten und die dieses mittels Infektion auf andere Zellen übertragen können. Ein Beispiel für ein solches Kon-
 strukt ist in Fig. 8 dargestellt. Es besteht auch die Möglichkeit, das hier beschriebene erfindungsgemäße HCV-
 Subgenom-Konstrukt mit integriertem Fremdgen direkt als Expressionsvektor einzusetzen. Dabei wird analog dem vor-
 genannten Verfahren vorgegangen, allerdings mit dem Unterschied, daß Zelllinien transfiziert werden, die keine trans-
 komplementierenden Faktoren exprimieren. In diesem Fall dient das HCV-Konstrukt also lediglich als
 Expressionsvektor.

Beispiel 7: Herstellung zellkultur-adaptierter HCV-RNA-Konstrukte

(A) Isolationsverfahren

[0078] Für die Bestimmung adaptiver Mutationen und die Herstellung zellkultur-adaptierter HCV-RNA-Konstrukte
 wurde wie folgt verfahren: Zellen wurden mit einem HCV-RNA-Konstrukt wie unter den Beispielen 1 und 2 beschrieben
 transfiziert und G418-resistente Zellklone hergestellt. Zur Bestimmung der Replikationskompetenz (darunter wird in
 diesem Zusammenhang die Anzahl G418-resistenter Zellklone verstanden, die pro Mikrogramm transfizierter HCV-
 RNA bzw. HCV-RNA-Konstrukt erhalten wird) wurde exemplarisch die Gesamt-RNA aus einem der Zellklone, genannt
 9-13 (Fig. 1B, Spur 11) isoliert und die Menge der darin enthaltenen HCV-RNA mittels Northern-blot wie in Fig. 2 B
 beschrieben bestimmt. 10 Mikrogramm der Gesamt-RNA, die ca. 10^9 Moleküle HCV-RNA enthielt, wurde anschließend
 per Elektroporation in naive Huh-7 Zellen eingeschleust (Fig. 9). Parallel dazu wurden 10^9 in vitro Transkripte der ana-
 logen *neo*-HCV-RNA, die mit isolierter Gesamt-RNA aus naiven Huh-7 Zellen auf eine Gesamt-RNA-Menge von 10 µg
 aufgefüllt worden war, in naive Huh-7 Zellen transfiziert. Nach Selektion mit G418 wurde die Anzahl der Zellkolonien,
 ausgedrückt in 'colony forming units (cfu) pro Mikrogramm RNA' in den beiden Ansätzen bestimmt. Bei einer Konzen-
 tration von 500 µg/ml G418 im Selektionsmedium betrug die Zahl der Kolonien, die mit der in der *isolierten* Gesamt-
 RNA enthaltenen HCV-RNA erhalten wurde, ca. 100.000 cfu pro Mikrogramm HCV-RNA. Dagegen wurden mit dersel-

ben Menge in vitro transkribierter HCV-RNA nur 30 - 50 Kolonien erhalten. Dieses Ergebnis belegt, daß die spezifische Infektiosität der HCV-RNA, die aus den Zellklonen isoliert wurde, ca. 1.000 - 10.000-fach höher ist als die Infektiosität der analogen in vitro Transkripte. Das methodische Vorgehen ist in Fig. 9 dargestellt.

[0079] Mit Hilfe der 'long-distance RT-PCR' wurde die HCV-RNA aus der Gesamt-RNA der 9-13 Zellen amplifiziert, die PCR-Amplifikate kloniert und zahlreiche Klone sequenziert. Ein Vergleich der Sequenzen dieser reklonierten RNAs mit der Sequenz der RNA, die ursprünglich in die naiven Huh-7 Zellen eingeschleust wurde ergab, daß die reklonierten RNAs zahlreiche Aminosäureaustausche besaßen, die über die gesamte HCV-Sequenz verteilt waren (Fig. 10). *Sfi*I-Fragmente dieser reklonierten Mutanten wurden im Austausch gegen das analoge *Sfi*I-Fragment des ursprünglichen Replikonstrukts in dieses eingeführt und RNAs der jeweiligen Mutanten wurden in naive Huh-7 Zellen eingeschleust. Nach Selektion mit G418 wurde dann für jede HCV-RNA-Mutante die Zahl der gebildeten Kolonien bestimmt. Während mit der Ausgangs-RNA nur 30 - 50 Kolonien pro Mikrogramm RNA erhalten wurde war die Koloniezahl bei zwei der reklonierten Varianten deutlich höher (Fig. 10). Im Fall der HCV-RNA-Konstrukte 9-13I und 9-13C betrug die spezifische Infektiosität 100 - 1.000 cfu pro Mikrogramm RNA und beim 9-13F Replikon sogar 1.000 - 10.000 cfu pro Mikrogramm RNA. Diese Ergebnisse zeigen, daß die Aminosäureaustausche in dem analysierten NS3-5B-Bereich der Mutanten 9-13I, 9-13C und insbesondere 9-13F zu einer deutlichen Erhöhung der Replikationskompetenz führten. Demgegenüber waren alle anderen HCV-RNA-Konstrukte (9-13 A, B, G, H und K) nicht mehr replikationskompetent, enthielten also letale Mutationen.

[0080] Zwecks Beantwortung der Frage, welche der Aminosäureaustausche im 9-13F-Konstrukt zur Steigerung der Replikation führten, wurden die Austausche einzeln oder in Kombination in das Ausgangs-HCV-RNA-Konstrukt eingeführt und die entsprechenden RNAs in naive Huh-7 Zellen eingeschleust. Das Ergebnis der Transfektionen mit diesen RNAs ist in Tabelle 1 zusammengefaßt. Daraus wird ersichtlich, daß im vorliegenden Beispiel die hohe Replikationskompetenz durch mehrere Mutationen bedingt ist. Den größten Beitrag leisten die Aminosäureaustausche in den HCV-RNA-Abschnitten NS5A und NS4B. Auch die einzelnen Austausche in der NS3-Region leisten einen Beitrag, der möglicherweise auf einem Synergismus dieser Einzelaustausche beruht.

Diese Befunde belegen, daß es durch die G418-Selektion der Zellen, die mit den *neo*-HCV-RNA-Konstrukten transfiziert wurden, zur Anreicherung solcher HCV-RNAs kam, die eine deutlich höhere Replikationskompetenz hatten. Mit dem hier beschriebenen Versuchsansatz lassen sich HCV-RNA-Konstrukte mit sehr unterschiedlicher Replikationseffizienz selektionieren. Je höher die Konzentration des Antibiotikums in dem Selektionsmedium ist, in/auf dem die HCV-RNA-Konstrukt-haltigen Zellen zwecks Selektion kultiviert werden, desto höher muß der Grad an adaptiven Mutationen und damit die Replikationseffizienz in den betreffenden HCV-RNA-Konstrukten sein, damit die Zellen auswachsen können. Werden die Selektionen mit niedrigeren Antibiotikum-Konzentrationen durchgeführt, können auch solche Zellen überleben und sich vermehren, die im Vergleich geringer adaptive Mutationen und eine weniger hohe Replikationseffizienz aufweisen.

Das bisher beschriebene HCV-RNA-Konstrukt 9-13F, das mehrere adaptive Mutationen enthielt, hatte eine erwiesenermaßen höhere Replikationseffizienz als die parentale HCV-RNA. Um HCV-RNAs mit noch höherer Replikation in Zellkultur zu erhalten, wurde die HCV-RNA, die in der Gesamt-RNA eines ausgewählten Zellklons enthalten war, mehrfach in naiven Huh-7 Zellen passagiert. Dieser ausgewählte Zellklon, genannt 5-15, wurde durch Transfektion mit dem HCV-RNA-Konstrukt I₃₈₉/NS3-3' erhalten (Fig. 1). Er entspricht weitgehend dem Zellklon 9-13, der durch Transfektion mit einem HCV-RNA-Konstrukt hergestellt wurde, das eine um 22 Nukleotide kürzere HCV-IRES besaß (I₃₇₇/NS3-3'; Fig. 1). 10 Mikrogramm Gesamt-RNA, isoliert aus dem Zellklon 5-15, wurden mittels Elektroporation in naive Huh-7 Zellen eingeschleust und die Zellen einer Selektion mit 1 mg/ml G418 unterzogen. Aus einem der so erzeugten Zellklone wurde wiederum Gesamt-RNA isoliert, in naive Huh-7 Zellen transfiziert und analog selektioniert. Dieser Vorgang wurde insgesamt viermal wiederholt. Nach der vierten Passage wurde aus einem Zellklon die Gesamt-RNA isoliert und die *neo*-HCV-RNA mit Hilfe der 'long-distance RT-PCR' amplifiziert. Das amplifizierte DNA-Fragment wurde mit dem Restriktionsenzym *Sfi*I verdaut und in das *Sfi*I-restringierte Ausgangskonstrukt I₃₈₉/NS3-3' inseriert. Insgesamt wurden über 100 DNA-Klone erhalten und zunächst mittels Restriktionsverdau analysiert. In vitro transkribierte RNA von ca. 80 dieser Klone wurde jeweils in naive Huh-7 eingeschleust und einer Selektion mit 500mg/ml G418 unterzogen. Von den 80 untersuchten *neo*-HCV-RNA-Varianten erwiesen sich die allermeisten als replikationsdefekt. Bei zwei Mutanten, genannt 5.1 und 19, war die spezifische Infektiosität, ausgedrückt als 'colony forming units' pro Mikrogramm RNA, jedoch sehr deutlich erhöht (Tabelle 2). Durch mehrfache Passage der RNA in Zellkultur lassen sich offensichtlich HCV-RNAs herstellen, deren Replikationseffizienz aufgrund von Mutationen (sog. "adaptiven Mutationen") mehrere Größenordnungen höher ist als die der ursprünglich aus dem Patienten klonierten RNA.

(B) Modifikationsverfahren

[0081] Solche nach (A) erzeugten und identifizierten adaptiven Mutationen können in ein wenig replikationskompetentes HCV-RNA-Konstrukt übertragen werden und führen zu einer massiven Steigerung der Replikation dieses Konstrukts. Diese Steigerung ist so hoch, daß damit nachweislich HCV-RNAs in Zellkultur zur Replikation gebracht

werden können, die kein selektierbares Markergen mehr besitzen. Fig. 12 zeigt einen Vergleich der Replikationseffizienz von HCV-RNAs, die entweder der Ausgangssequenz oder den adaptierten Sequenzen 9-13F bzw. 5.1 entsprechen. Zwecks einfacher Messung wurde das *neo*-Gen entfernt und durch das Gen für die Luziferase ersetzt. Als Negativkontrolle diente wiederum ein HCV-RNA-Konstrukt, das auf Grund einer inaktivierenden Mutation der NS5B RNA-Polymerase replikationsdefekt war. Schon 24 Stunden nach der Transfektion erkennt man einen deutlichen Unterschied in der Luziferaseaktivität zwischen der defekten RNA und den 9-13F bzw. 5.1-Konstrukten während zwischen der defekten RNA (318 DN) und dem Ausgangs-RNA-Konstrukt (wt) das keine adaptiven Mutationen besaß, kaum ein Unterschied zu sehen war. Während des gesamten Beobachtungszeitraums wurde die höchste Luziferaseaktivität und damit die höchste Replikation mit der 5.1-RNA erhalten. Diese Befunde belegen nicht nur die hohe Replikationseffizienz dieser RNA, sondern zeigen auch, daß es möglich ist, mit adaptierten HCV-RNA-Konstrukten ein Zellkultursystem aufzubauen, für das die Anwesenheit eines selektierbaren Gens nicht mehr notwendig ist. Eine zusammenfassende Übersicht der Nukleotid- und Aminosäureunterschiede zwischen dem Ausgangskonstrukt und den Mutanten 9-13F, 5.1 und 19 ist in Tabelle 3 gegeben.

15 **Beispiel 8: Herstellung zellkultur-adaptierter HCV-RNA-Vollängengenome**

[0082] In den Beispielen 1 bis 7 wurde stets eine subgenomische HCV-RNA verwendet, der die gesamte Strukturproteinregion von Core bis einschließlich p7 bzw. NS2 fehlte. Im vorliegenden Beispiel 8 wird gezeigt, daß es möglich ist, mit Hilfe der adaptierten NS3-5B-Sequenz ein HCV-Vollängengenom in Zellkultur zur Replikation zu bringen. Zu diesem Zweck wurde zunächst das Sfil-Fragment der gemäß Beispiel 7 hergestellten, hoch adaptierten HCV-RNA 5.1 in ein selektionierbares HCV-Vollängengenom transferiert (Fig. 12). Dieses HCV-Genom wurde in naive Huh-7 Zellen transfiziert und einer Selektion mit unterschiedlichen G418-Konzentrationen unterzogen. In Abhängigkeit von der Selektionsstärke (der G418-Konzentration) wurde eine unterschiedlich große Zahl an Zellklonen erhalten (Fig. 12 B). Im Vergleich dazu wurden mit dem unveränderten HCV-Vollängengenom, das keine adaptiven Mutationen enthielt, keine Kolonien erhalten, ebenso mit der Negativkontrolle, die auf Grund einer inaktivierenden Mutation in der NS5B RNA-Polymerase replikationsdefekt war. Zum Nachweis dafür, daß die so entstandenen Zellklone tatsächlich ein autonom replizierendes HCV-Vollängengenom enthielten, wurde Gesamt-RNA aus mehreren Zellklonen isoliert und mittels Northern-Blot analysiert. In allen Zellklonen war die Vollängen HCV-RNA eindeutig nachweisbar (Fig. 12). Damit ist eindeutig belegt, daß es mit Hilfe der an Zellkulturen adaptierten HCV-Sequenzen möglich ist, ein HCV-Vollängengenom herzustellen, das mit hoher Effizienz und autonom in einer Zelllinie repliziert, d.h. es können mit dem erfindungsgemäßen System auch adaptierte HCV-Vollängengenome hergestellt werden. Da dieser Klon darüber hinaus die vollständige HCV-Sequenz besitzt, also auch die für die Viruspartikelbildung notwendigen Strukturproteine, ist es mit diesem System möglich, große Mengen infektiöser Viruspartikel in Zellkulturen herzustellen. Zum Nachweis dieser Viren werden zellfreie Überstände von Zellen, die ein replizierendes HCV-Vollängengenom tragen, auf naive Huh-7 Zellen gegeben und die so infizierten Zellen einer Selektion mit G418 unterzogen. Jeder Zellklon, der unter diesen Bedingungen auswächst, geht auf eine infizierte Zelle zurück. Die Viren in den Zellkulturüberständen von Zellen, die ein replizierendes HCV-Vollängengenom besitzen, können aber auch mit verschiedenen im Stand der Technik bekannten Verfahren wie Ultrazentrifugation oder Mikrodialyse angereichert und gereinigt werden und dann zur Infektion naiver Zellen verwendet werden. Mit diesem Verfahren ist eindeutig gezeigt, daß mit dem erfindungsgemäßen HCV-Zellkultursystem zellkultur-adaptierte HCV-Vollängengenome hergestellt werden können, die mit hoher Effizienz in Zellen replizieren und infektiöse Viren produzieren. Diese können ebenfalls durch Infektion eines Versuchstiers, vorzugsweise dem Schimpansen, nachgewiesen werden.

45 **Beispiel 9: Herstellung von HCV-Vollängen-Konstrukten und HCV-Subgenom-Konstrukten mit Reportergen.**

[0083] Es wird ein HCV-RNA-Konstrukt hergestellt, bei dem anstelle des Antibiotikumresistenzgens ein Reportergen eingefügt wird (Fig. 13). Dabei kann die Replikation anhand der Menge bzw. der Aktivität des Reportergens bzw. Reportergenprodukts bestimmt werden. Das Reportergen ist vorzugsweise ein Gen aus der Gruppe der Luziferasegene, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), dem GFP-Gen (green fluorescence protein Gen), dem GUS-Gen (Glukuronidasegen) oder dem SEAP-Gen (sezernierte alkalische Phosphatasegen). Diese Reportergene bzw. deren Produkte, nämlich die entsprechenden Reporterproteine, können z.B. mittels Fluoreszenz, Chemilumineszenz, colorimetrisch oder mit Hilfe immunologischer Methoden (z.B. enzyme-linked immunosorbent assay, ELISA) bestimmt werden. Das Reportergen kann entweder von einer eigenen IRES exprimiert werden oder in Form eines Fusionsproteins, das entweder als solches aktiv ist oder mittels einer proteolytisch spaltbaren Aminosäuresequenz so mit einem HCV-Protein verbunden ist, daß es von einer zellulären oder viralen (HCV-)Protease von diesem abgespalten wird.

Beispiel 10: Herstellung von HCV-Vollängen-Konstrukten mit integrierten Fremdgenen zur Verwendung als leberzellspezifische Genfährer für die Gentherapie oder als Expressionsvektor.

[0084] Das Konstrukt (Fig. 14) wird in Zellen eingeschleust und führt dort zur Bildung von HCV-Viruspartikeln, die zur Infektion weiterer Zellen verwendet werden können. Da die Viruspartikel eine RNA mit einem Fremdgen enkapsidiert haben kann dieses in den so infizierten Zellen zur Produktion des von diesem Fremdgen kodierten Proteins benutzt werden. Zellen, die mit dem Konstrukt transfiziert wurden, exprimieren ebenfalls das Fremdgen.

Beispiel 11: Herstellung von monocistronischen HCV-RNA-Konstrukten, bei denen das Resistenzgenprodukt als Fusionsprotein mit dem HCV-Anteil exprimiert wird.

[0085] Für bestimmte Untersuchungen ist es von Vorteil, wenn das HCV-RNA-Konstrukt kein heterologes IRES-Element besitzt. Solche Untersuchungen sind beispielsweise die Bestimmung der Interferonresistenz. Wird eine Zelle, die ein HCV-RNA-Konstrukt besitzt, mit Interferon-alpha oder -beta inkubiert, kommt es zu einer Reduktion der Replikation der HCV-RNA. Zur Aufklärung des Wirkungsmechanismus ist es notwendig, daß das HCV-RNA-Konstrukt keine heterologe IRES besitzt, da ansonsten nicht bestimmt werden kann, ob die Interferon-vermittelte Hemmung durch eine Hemmung der HCV-Replikation oder durch eine Hemmung der heterologen IRES vermittelt wird. Deshalb werden Konstrukte hergestellt, bei denen das Resistenzgen mit einem HCV-Protein fusioniert wird (Fig. 15). Entweder das Fusionsprotein ist als solches aktiv oder das Resistenzgenprodukt wird mittels einer proteolytisch spaltbaren Aminosäuresequenz so mit einem HCV-Protein verbunden ist, daß es von einer zellulären oder viralen (HCV-)Protease von diesem abgespalten wird.

Tabelle 1

Spezifische Infektiositäten (cfu/µg RNA) der HCV-RNA-Konstrukte mit adaptiven Mutationen, die bei der 9-13F Mutante gefunden und in das parentale HCV-RNA-Konstrukt I ₃₈₉ /NS3-3'/wt eingeführt wurden		
Aminosäureaustausch ¹	HCV-Protein	cfu/µg RNA ²
kein		30 - 60
1283 arg -> gly	NS3	200 - 250
1383 glu -> ala	NS3	30 - 60
1577 lys -> arg	NS3	30 - 60
1609 lys -> glu	NS3	160 - 300
(1283 arg -> gly + 1383 glu -> ala + 1577 lys -> arg + 1609 lys -> glu)	NS3	360 - 420
1936 pro -> ser	NS4B	500 - 1000
2163 glu -> gly	NS5A	1000-5000
2330 lys -> glu	NS5A	30 - 60
2442 ile -> val	NS5B	30 - 60
alle zusammen		5000

¹ Aminosäureaustausch im Polyprotein des HCV-Isolats Con-1 (EMBL-Genbank No. AJ238799); die Aminosäuren sind im Dreibuchstabenkode angegeben.

² Colony forming units (Anzahl der Zellklone) bei einer Selektion von 500µg/ml G418.

Tabelle 2

Spezifische Infektiositäten (cfu/μg RNA) des parentalen HCV-RNA-Konstrukts I ₃₈₉ /NS3-3'/wt und der Varianten 9-13C, 9-13I, 9-13F, 5.1 und 19.		
Transfizierte RNA-Variante	cfu/μg RNA ¹	
Wildtyp	30 - 50	
9-13 C	100 - 1.000	
9-13 I	100 - 1.000	
9-13 F	1.000 - 10.000	
5.1	50.000 - 100.000	
19	50.000 - 100.000	

¹ Colony forming units (Anzahl der Zellklone) bei einer Selektion von 500μg/ml G418.

**Tabelle 3: Nukleotid- und Aminosäureunterschiede zwischen dem parental HCV-
RNA-Konstrukt I₃₈₉/NS3-3'/wt und den Mutanten 9-13I, 9-13F, 5.1 und 19**

HCV Mutante	Nukleotidposition	Nukleotidaustausch	Aminosäureaustausch
9-13 I	3685	C > T	Pro > Leu
	4933	C > T	Thr > Met
	5249	T > C	-
	8486	C > T	-
	8821	G > A	Trp > stop
	8991	C > G	Arg > Gly
	9203	A > G	-
	9313	T > C	Phe > Ser
	9346	T > C	Val > Ala
9-13 F	3866	C > T	-
	4188	A > G	Arg > Gly
	4489	A > C	Glu > Ala
	4562	G > A	-
	4983	T > C	-
	5071	A > G	Lys > Arg
	5166	A > G	Lys > Glu
	6147	C > T	Pro > Ser
	6829	A > G	Glu > Gly
	7329	A > G	Lys > Glu
	7664	A > G	Ile > Val
	8486	C > T	-
	8991	C > G	Arg > Gly
NK5.1	4180	C > T	Thr > Ile
	4679	C > T	-

5		4682	T > C	-
		5610	C > A	Leu > Ile
		6437	A > G	-
10		6666	A > G	Asn > Asp
		6842	C > T	-
		6926	C > T	-
15		6930	T > C	Ser > Pro
		7320	C > T	Pro > Ser
		7389	A > G	Lys > Glu
20	NK19	3946	A > G	Glu > Gly
		4078	C > G	Ala > Gly
		4180	C > T	Thr > Ile
25		4682	T > C	-
		5610	C > A	Leu > Ile
		5958	A > T	Met > Leu
30		6170	T > A	-
		6596	G > A	-
35		6598	C > G	Ala > Gly
		6833	C > T	-
		6842	C > T	-
40		6930	T > C	Ser > Pro
		7141	A > G	Glu > Gly
		7320	C > T	Pro > Ser
45		7389	A > G	Lys > Glu
		7735	G > A	Ser > Asn
50				

[0086] Angegeben sind die Unterschiede der Nukleotid- und Aminosäuresequenzen zwischen der Ausgangs-HCV-RNA-Sequenz Con 1 (EMBL-Genbank No. AJ238799) und denen der zellkulturadaptierten HCV-RNAs. Die Zahlen
55 beziehen sich auf die Nukleotid- und Aminosäurepositionen des HCV-Isolats Con1.

EP 1 043 399 A2

SEQUENCE LISTING

5 <110> Bartenschlager, Ralf

<120> Hepatitis C Virus cell culture system

<130> ba-1

<140> 199 15 178.4

10 <141> 1999-04-03

<160> 11

<170> PatentIn Ver. 2.1

15 <210> 1

<211> 11076

<212> DNA

<213> Hepatitis C Virus

<400> 1

20 gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60

tctttcacgca gaaagcgtct agccatggcg ttagtatgag tgctcgtgcag cctccaggac 120

ccccctcccc gggagagcca tagtggctcg cggaaccggt gactacaccg gaattgccag 180

gacgacccggg tcttttcttg gatcaacccg ctcaatgcct ggagatttgg gcgtgccccc 240

gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300

gtgcttgcga gtgccccggg aggtctcgtg gaccgtgcac catgagcacg aatcctaacc 360

ctcaaaagaaa aaccaaaccg aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420

acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480

25 caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 540

ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600

cgtggctggc caccgacggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660

gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720

ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780

cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840

30 tggaaagccgg tcttgtcgat caggatgacg tggacgaaga gcacagggg ctcgcgccag 900

ccgaactgtt cgccaggctc aaggcgcgca tgcccagcgg cgaggatctc gtcgtgacct 960

atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020

actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080

ttgctgaaga gcttggcggc gaattggctg accgcttctt cgtgctttac ggtatcgccg 1140

ctcccgatcc gcagcgcacg gccttctatc gccttcttga cgagtctctc tgagtttaaa 1200

cagaccacaa cgggttccct ctgacgggat caattccgcc cctctccctc cccccccctt 1260

35 aacgttactg gccgaagccg tctggaataa ggccggtgtg cgtttgtcta tatgttattt 1320

tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380

acgagcattc ctgagggtct tccccctctc gccaaaggaa tgcaaggctc gttgaatgtc 1440

gtgaagggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500

tgacggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560

taagatacac ctgcaaaagg gccacaaccg cagtgccacg ttgtgagttg gatagttgtg 1620

40 gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680

gtaccccatg gtatgggacg tgatctgggg cctcgttgca catgctttac atgtgtttag 1740

tcgagggttaa aaaacgtcta ggcccccgga accacgggga cgtggttttc ctttgaaaaa 1800

cacgataata ccatgggcac gaatcctaaa cctcaaagaa aaaccaaacy taacaccaac 1860

cgccgcccac aggacgtcaa gttcccgggc ggtggtcaga tcgtcgggtg agtttacctg 1920

ttgccgcgca ggggccccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggctg 1980

45 caacctcgtg gaaggcgaca acctatcccc aaggctcgcc agcccagagg tagggcctgg 2040

gctcagcccg ggtacccctg gcccctctat ggcaatgagg gcttgggggtg ggcaggatgg 2100

ctcctgtcac cccgtggctc tcggcctagt tggggcccca cgacccccg gcgtaggtcg 2160

cgcaatttgg gtaagggtcat cgataccctc acgtgcggct tcgccgatct catggggtac 2220

attccgctcg tcggcgcccc cctagggggc gctgccaggg ccctggcgca tggcgtccgg 2280

gttctggagg acggcggtgaa ctatgcaaca gggaatctgc ccggttgcct cttttctatc 2340

50 ttccttttgg ctttgcgtgc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400

gtatccggag tgtaccatgt caccgaacgac tgctccaacg caagcattgt gtaggagca 2460

gcggacatga tcatgcatac ccccggtgac gtgccctgcg ttcgggagaa caactcctcc 2520

cgctgctggg tagcgtcac tcccacgctc gcggccagga acgctagcgt ccccaactacg 2580

acgatacgac gccatgtcga tttgctcgtt gggggcggctg ctctctgctc cgctatgtac 2640

gtggggagatc tctgcggatc tgttttctc gtcgcccagc tgttcaacct etcgcctcgc 2700

EP 1 043 399 A2

5 cggcacgaga cagtacagga ctgcaattgc tcaatatatc ccggccacgt gacagggtcac 2760
 cgtatggcctt gggatatgat gatgaactgg tcacctacag cagccctagt gggtatcgag 2820
 ttactccgga tcccacaagc tgtcgtggat atggtggcgg gggccattg gggagtccta 2880
 ggcggcccttg cctactatct catggtgggg aactgggcta aggttctgat tgtgatgcta 2940
 ctctttgccc gcgttgacgg gggaaacctat gtgacagggg gacgatggc caaaaacacc 3000
 ctccgggatta cgtccctctt ttcacccggg tcatccaga aaatccagct tgtaaacacc 3060
 aacggcagct ggcacatcaa caggactgcc ctgaactgca atgactccct caacactggg 3120
 ttcccttgctg cgctgttcta cgtgcacaag ttcaactcat ctggatgccc agagcgcatg 3180
 gccagctgca gcccacatga cgcgttcgct cagggggtgg ggcccatcac ttacaatgag 3240
 10 tcacacagct cggaccagag gccttattgt tggcactacg caccocggcc gtgcggtatc 3300
 gtaccocggg cgcagggtgt tggccagatg tactgcttca ccccaagccc tgtcgtgggtg 3360
 gggacgagcg accggttcgg cgtccctacg tacagttggg gggagaatga gacggacgtg 3420
 ctgcttctta acaacacggc gccgcccga ggcaactggg ttggctgtac atggatgaat 3480
 agcactgggt tcaaccaagc gtgcgggggc ccccctgta acatcggggg gatcggcaat 3540
 aaaacctga cctgccccac ggaactgttc cggaaagcacc ccgaggccac ttacaccaa 3600
 15 tgtggttcgg ggcttgggtt gacacccaga tgcttgggtcc actacccata caggctttgg 3660
 cactacccct gcaactgtcaa cttaaccatc ttcaagggtta ggaatgacgt ggggggagtg 3720
 gagcacaggg tcgaagccgc actcaattgg actcagaggag agcgttgtaa cctggaggac 3780
 agggacagat cagagcttag cccgctgctg ctgtctacaa cggagtggca ggtattgccc 3840
 tgttctctca ccacctacc ggctctgtcc actgggttga tccatctcca tcagaacgtc 3900
 gtggacgtac aatacctgta cggatatagg tcggcgggtg tctcctttgc aatcaaatgg 3960
 20 gagtatgtcc tgttgcctct ccttctctct gcgacgcgc gcgtctgtgc ctgcttggg 4020
 atgatgctgc tgatagctca agctgaggcc gccctagaga acctgggtgt cctcaacggc 4080
 gcatccgttg ccggggcgca tggcaattct tccctcctcg tgttctctg tgcgtcctg 4140
 tacatcaagg gcaggctggg cctggggcg gcatatgccc tctaaggcgt atggccgcta 4200
 ctccctgctcc tgctggcggt accaccacga gcatacgcca tggaccggga gatggcagca 4260
 tcgtgcggag gcgcgggttt cgtaggctcg atactcttga ccttgtcacc gactataag 4320
 ctgttctctg ctaggctcat atggtgggtta caatatttta tcaccagggc cgaggcacac 4380
 ttgcaagtgt ggatcccccc cctcaacgtt cggggggggc gcgatgcccgt catcctcctc 4440
 25 acgtgcgcga tccacccaga gctaactctt accatcacca aaatcttctg ctgacttctg gcgcgcacac 4500
 ggtccactca tgggtgctcca ggctgggtata accaaagtgc cgtacttctg gcgcgcacac 4560
 gggctcattc gtgcattgat gctgggtgcg aaggttgcgt ggggtcatta tgtccaaatg 4620
 gctctcatga agttggccgc acgtacgttt atgacctct caccacctg 4680
 cgggactggg cccacgcggg cctacgagac ctgtcgggtg cagttagacc cgtcgtcttc 4740
 tctgatattg agaccaaggt tateacctgg ggggcagaca ccgcgcggtg tggggacatc 4800
 30 atcttggggc tgcccgtctc cgcgcgaggg gggaggagga tacatctggg accggcagac 4860
 agccttgaag ggcagggggt gcgactcctc gcgcctatta cggcctactc ccaacagacg 4920
 cgaggcctac ttggctgcat ctacactagc ctacaggcc gggacaggaa ccaggctcag 4980
 ggggaggttc aagtgggtct caccgcaaca caatcttctc tggcgacctg cgtcaatggc 5040
 gtgtgttggg ctgtctatca tgggtccggc tcaaagaccc ttgccggccc aaaggggcca 5100
 atcacccaaa tgtacaccaa tgtggaccag gacctcgtcg gctggcaagc gcccccggg 5160
 35 gcgcgttctt tgacaccatg cactctgcgc agctcggacc tttacttggg cagcaggcat 5220
 gccgatgtca ttccgggtgc cggcgggggc gacagcaggg ggagcctact ctccccagg 5280
 cccgtctcct acttgaaggg ctcttcgggc ggtccactgc tctgcccctc ggggcacgct 5340
 gtgggcatct ttccgggtgc cgtgtgcacc cgaggggttg cgaaggcggt ggaacttcta 5400
 cccgtcgagt ctatggaaac cactatgcgg tccccggtct tcacggacaa ctgcgtccct 5460
 ccggccgtac cgagacatt ccagggtggc catctacacg cccctactgg tagcggaag 5520
 40 agcactaagg tgccggctgc gtatgcagcc caagggata aggtgcttgt cctgaacccg 5580
 tccgtcgccg ccacctagg tttcggggcg tatatgtcta aggcacatgg tatcgacct 5640
 aacatcagaa ccggggtaag gacctacac acgggtgccc ccatcacgta ctccacctat 5700
 ggcaagtctt ttgcccagcg tgggtgtctc ggggggcctt atgacatcat aatatgtgat 5760
 gagtgcact caactgactc gaccactatc ctgggcatcg gcacagtcct ggaccaagcg 5820
 gagacggctg gagcgcgact cgtcgtgctc gccaccgcta cgcctccggg atcggtcacc 5880
 gtgccacatc caaacatcga ggaggtggct ctgtccagca ctggagaaat ccccttttat 5940
 45 ggcaagcca tccccatcga gaccatcaa gggtggaggc acctcatttt ctgccattcc 6000
 aagaagaaat gtgatgagct cgcgcgaa cgttccggcc tcggactcaa tgcgttagca 6060
 tattaccggg gccttgatgt atccgtcata ccaactagcg gagacgtcat tgtcgtagca 6120
 acggacgctc taatgacggg ctttaccggc gatattcgact cagtgatcga ctgcaatata 6180
 tgtgtcacc agacagtcga cttcagcctg gacccgacct tcaccattga gacgacgacc 6240
 gtgccacaag acgcgggtgc acgctcgag cggcgaggca ggactggtag gggcaggatg 6300
 50 ggcatttaca ggtttgtgac tccaggagaa cggccctcgg gcatgttcga ttcctcgggt 6360
 ctgtgcgagt gctatgacgc gggctgtgct tggtagagc tcacgcccgc cgagacctca 6420
 gttagggttc gggcttacct aaacacacca gggttgccc tctgccagga ccatctggag 6480
 ttctgggaga gcgtctttac aggcctcacc cacatagacg cccatttctt gtcccagat 6540
 aagcaggcag gagacaactt cccctacctg gtacatacc aggctacggg gtgcgcccag 6600
 gctcaggctc cacctccatc gtgggaccaa atgtggaagt gtctcatagc gctaaagcct 6660

55

EP 1 043 399 A2

	acgctgcacg	ggccaacgcc	cctgctgtat	aggctgggag	ccgttcaaaa	cgaggttact	6720
	accacacacc	ccataaccaa	atcacatcatg	gcatgcatgt	cggttgacct	ggaggtcgct	6780
	acgagccact	gggtgctggg	aggcggagtc	ctagcagctc	tggccgcgta	ttgcctgaca	6840
5	acagggcagc	tggtcattgt	gggcaggatc	atcttgtccg	gaaagccggc	catcattccc	6900
	gacaggggag	tcctttaccg	ggagttcgat	gagatggaag	agtgccgctc	acacctccct	6960
	tacatcgaa	aggggaatgca	gctcgccgaa	caattcaaac	agaaggcaat	cggttgctg	7020
	caaacagcca	ccaagcaagc	ggaggctgct	gctcccgtgg	tggaatccaa	gtggcggacc	7080
	ctcgaagcct	tctgggagaa	gcatatgtgg	aatttcatca	gcgggataca	atatttagca	7140
	ggcttgctca	ctctgcctgg	caaccccgcg	atagcatcac	tgatggcatt	cacagcctct	7200
	atcaccagcc	cgctcaccac	ccaacatacc	ctcctgttta	acatcctggg	gggatgggtg	7260
10	gccgcccac	ttgctcctcc	cagcgctgct	tctgtcttcg	taggcgcggg	catcgctgga	7320
	gcggctgttg	gcagcatagg	ccttgggaag	gtgcttggtg	atattttggc	aggttatgga	7380
	gcaggggtgg	cagggcgctg	cgtggccttt	aagggtcatga	gcggcgagat	gccctccacc	7440
	gagggacctg	ttaacctact	ccctgtctatc	ctctcccctg	gcgccttagt	cgctgggggtc	7500
	gtgtgcctgg	cgatactgcg	tcggcacgtg	ggccacgggg	agggggctgt	gcagtggatg	7560
	aaccggctga	tagcgttcgc	ttcgcggggt	aaccacgtct	ccccacgca	ctatgtgcct	7620
15	gagagcgacg	ctgcagcagc	tgctcactcag	atcctctcta	gtcttaccat	cactcagctg	7680
	ctgaagaggc	ttcaccagtg	gactgctcca	cgccatgctc	cggtcggtgg	7740	
	ctaagagatg	tttgggattg	gatatgcacg	gtgttgactg	atctcaagac	ctggctccag	7800
	tccaagctcc	tgccgcgatt	gcggggagtc	cccttcttct	catgtcaacg	tggtgtacaag	7860
	ggagtctggc	ggggcgacgg	catcatgcaa	accacctgcc	catgtggagc	acagatcacc	7920
	ggacatgtga	aaaaacggtc	catgaggatc	gtggggccta	ggacctgtag	taaacacgtg	7980
20	catggaacat	tccccattaa	cgcgtacacc	acggggccct	gcacgcccct	cccggcgcca	8040
	aattattcta	ggcgctgtgt	gctgaggagt	acgtggaggt	acgtggaggt	tacgggggtg	8100
	ggggatttcc	actacgtgac	gggcatgacc	actgacaacg	taaagtgcct	gtgtcagggt	8160
	ccggcccccg	aattctctac	agaagtggat	gggtgtcggt	tgacagagta	cgctccagcg	8220
	tgcaaacccc	tcttacggga	ggaggtcaca	ttcctggctg	ggctcaatca	atacctggtt	8280
	gggtcacagc	tcccattgca	gcccgaaccc	gacgtagcag	tgctcacttc	catgtccacc	8340
	gacccctccc	acattacggc	ggagacggct	aagcgtaggc	tgccacgggg	atctcccccc	8400
25	tctttggcca	gctcatcagc	tagccagctg	tctgocgctt	ccttgaagcg	aacatgcact	8460
	acccgctcatg	actccccgga	cgctgacctc	atcgaggcca	acctcctgtg	gcggcaggag	8520
	atgggcgggg	acatcacccc	cgtggagtcg	gaaaataagg	tagtaatttt	ggactctttc	8580
	gagccgctcc	aagcggagga	ggatgagagg	gaagtatccg	ttccggcgga	gatcctgcga	8640
	aggtccagga	aattccctcg	agcgtatccc	atatgggcac	gcccggatta	caacccctcca	8700
	ctgttagagt	cctggaagga	ccgggactac	gtccctccag	tggtacacgg	gtgtccattg	8760
30	cgccctgcca	aggccctccc	gataccacct	ccacggagga	agaggacggt	tgctctgtca	8820
	gaatctaccg	tgtcttctgc	cttggcgag	ctcgccacaa	agaccttcgg	cagctccgaa	8880
	tcgtcgcccg	tcgacagcgg	cacggcaacg	gcctctcctg	accagccctc	cgacgacggc	8940
	gacgcgggat	ccgacgttga	gtcgtactcc	tccatgcccc	cccttgaggg	ggagccgggg	9000
	gctcccgatc	tcagcgacgg	gtcttggtct	accgtaaagc	aggaggctag	tgaggacgtc	9060
	gtctgtgctc	cgatgtccta	cacatggaca	ggcgccctga	tcacgccaatg	cgctggcgga	9120
	gaaaccaagc	tgcccatcaa	tgactgagc	aactctttgc	tcctgtacca	caacttggtc	9180
35	tagtctacaa	catctcgag	cgcaagcctg	cggaagaaga	aggtcacctt	tgacagactg	9240
	caggtccttg	acgaccacta	ccgggacgtg	ctcaaggaga	tgaaggcgaa	ggcgtccaca	9300
	gttaaggcta	aacttctatc	cgtggaggaa	gcctgtaagc	tgacgcccc	acattcggcc	9360
	agatctaaat	ttggctatgg	ggcaaaaggac	gtccggaaac	tatccagcaa	ggccgttaac	9420
	cacatccgct	ccgtgtggaa	ggacttgctg	gaagacactg	agacaccaat	tgacaccacc	9480
	atcatggcaa	aaaatgaggt	tttctgcgtc	caaccagaga	agggggggccg	caagccagct	9540
40	cgcccttatc	tattcccaga	tttgggggtt	cgtgtgtcg	agaaaatggc	cctttacgat	9600
	gtggctccca	ccctccctca	ggcctgtatg	ggctcttcat	acggattcca	atactctcct	9660
	ggacagcggg	tcgagttcct	gggtgaatg	tggaagcgga	agaaatgccc	tatgggcttc	9720
	gcatatgaca	cccgtgtttt	tgactcaacg	gtcactgaga	atgacatccg	tggtgaggag	9780
	tcaatctacc	aatgttgtga	cttggccccc	gaagccagac	aggccataag	gtcgtccaca	9840
	gagcggcttt	acatcggggg	ccccctgact	aattctaaag	ggcagaactg	cggtatctgc	9900
	cggtgcggcg	cgagcgggtg	actgacgacc	agctgcggta	ataccctcac	atgttacttg	9960
45	aaggccgctg	cgccctgtcg	agctgcgaag	ctccaggact	gcacgatgct	cgatggcgga	10020
	gacgaccttg	tcgttatctg	tgaaagcgcg	gggacccaag	aggacgaggg	gagcctacgg	10080
	gccttcacgg	aggctatgac	tagatactct	gccccccctg	gggacccggc	caaacagaa	10140
	tacgacttgg	agttgataac	atcatgctcc	tccaatgtgt	cagtcgcgca	cgatgcactt	10200
	ggcaaaaagg	tgtactatct	cacccgtgac	cccaccaccc	cccttgcgcg	qgctgcgtgg	10260
	gagacagcta	gacacactcc	agtcaattcc	tggttaggca	acatcatcat	gtatgcggcc	10320
	accttgtggg	caaggatgat	cctgatgact	catttcttct	ccatccttct	agctcaggaa	10380
50	caacttgaaa	aagccctaga	ttgtcagatc	tacggggcct	gttactccat	tgagccactt	10440
	gacctacctc	agatcattca	acgactccat	ggccttagcg	catttttact	ccatagttac	10500
	tctccagggt	agatcaatag	ggtggcttca	tgccctcagg	aacttggggg	accgcccctg	10560
	cgagtctgga	gacatcgggc	cagaagtgtc	cgcgctaggc	tactgtccca	ggggggggag	10620

EP 1 043 399 A2

5 gctgccactt gtggcaagta cctcttcaac tgggcagtaa ggaccaagct caaactcact 10680
 ccaatcccgg ctgcgtecca gttggattta tccagctggg tcgttgctgg ttacagcggg 10740
 ggagacatat atcacagcct gtctcgtgcc cgaccccgtt ggttcatgtg gtgcctactc 10800
 ctactttctg taggggtagg catctatcta ctccccaacc gatgaacggg gagctaaaca 10860
 ctccaggcca ataggccatc ctgttttttt cccttttttt ttttcttttt tttttttttt 10920
 tttttttttt tttttttttt tccttttttt ttctcttttt tttccttttc tttccttttg 10980
 tggctccatc ttageccctag tcacggctag ctgtgaaagg tccgtgagcc gcttgactgc 11040
 agagagtgct gatactggcc tctctgcaga tcaagt 11076

10

15

20

25

30

35

40

45

50

55

EP 1 043 399 A2

<210> 2
<211> 8637
<212> DNA
<213> Hepatitis C Virus

5

```

<400> 2
gccagccccc gattgggggc gacactccac catagatcac tccccgtgga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggtctg cggaaccggg gagtacaccg gaattgccag 180
gacgacccgg tcccttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagcccgagta gtgttgggtc gcgaaaggcc ttgtggtagt gcctgatagg 300
gtgcttgaga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatccataac 360
ctcaaaagaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc 420
cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct 480
ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgcgc ggttcttttt gtcaagaccg 540
acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggtatcgc tggctggcca 600
cgacggcggt tccctgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc 660
tgctattggg cgaagtgcgc gggcaggatc tccgtgcatc tcaccttgct cctgccgaga 720
aagatcccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc 780
cattcgacca ccaagcgaaa catcgcatcg agcagcacg tactcggatg gaagccggct 840
ttgtcgatca ggatgatctg gacgaagagc atcagggggt cgcgccagcc gaactgttcg 900
ccaggctcaa ggccgcgatg cccgacggcg aggatctcgt cgtgacccat ggcgatgcct 960
gcttgccgaa tatcatgggt gaaaaatggc gcttttcttg attcatcgac tgtggccggc 1020
tgggtgtggc ggaccgctat caggacatag cgttggttac cctgatatt gctgaagagc 1080
ttggcggcga atgggctgac cgcttctcgc tgctttacgg tatcgccgct ccgattctgc 1140
agcgcatcgc ctctctatcg ctctctgacg agttcttctg agtttaaaac gaccacaacg 1200
gttccctctc agcgggatca attccgcccc tctccctccc cccccctaa cgttactggc 1260
cgaagccgct tggaaataag ccggtgtgctg tttgtctata tgttattttc caccatattg 1320
ccgtcttttg gcaatgtgag gggccggaaa cctggccctg tcttcttgac gagcattcct 1380
aggggtcttt aggtctctgc caaaggaatg caaggtctgt tgaatgtcgt gaaggaagca 1440
gttctcttgg aagcttcttg aagacaaaca acgtctgtag cgacctttg caggcagcgg 1500
aaccacccac ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata agatcacact 1560
gcaaaaggcg cacaaccccc gtgccacggt gtgagttgga tagttgtgga aagatcacaa 1620
tggtctctct caagcgtatt caacaagggg ctgaaggatg ccagaagggt acccatttgt 1680
atgggatctg atctggggcc tcggtgcaca tgctttacat gtgtttagtc gaggttaaaa 1740
aacgtctatg aacgtctatg ccccccgaac cagggggacg tggtaaaaca cgataatacc 1800
atggaccggg agatggcagc atcgtgcgga ggcgcgggtt tcgtaggtct gatactcttg 1860
accttctcac cgcactataa gctgttccct gctaggctca tatggtggtt acaatatttt 1920
atcacagggc cagaggcaca ctltcaagtg tggatcccc cctcaacgt tcgggggggc 1980
cgcgatgcgc tcatctctct cacgtgcgcg atccacccag agctaattct taccatcacc 2040
aaaatcttgc tcgccatact cgggtccactc atgtgtctcc aggtcgttat aaccaaagt 2100
cgtactctcg tgccgcgaca cgtgctcatt cgtgcatgca tgctgtgctg gaaggttct 2160
gggggtcatt atgtccaaat ggctctcatg aagtgtggcg cactgacagg tacgtacgtt 2220
tatgaccatc tcaccccaact gcgggactgg gcccacgcgg gcctacgaga ccttgcgggt 2280
gcagttaggc ccgtcgtctt ctctgatatg gagaccaagg ttatcacctg gggggcagac 2340
accgcggcgt gtggggacat catcttgggc ctgcccgtct ccgcccgag ggggagggag 2400
atacatctgg gaccggcaga cagccttgaa gggcaggggt ggcgactcct cgcgcctatt 2460
acggcctact cccaacagac gcgaggccta cttggctgca tcatcactag cctcacagge 2520
cgggacagga accaggtcga gggggagggt caagtgtctt ccaccgcaac acaatctttc 2580
ctggcgacct gcgtcaatgg cgtgtgttgg actgtctatc atgtgtccgg ctcaaagacc 2640
cttgccggcc caaaggggcc aatcacccaa atgtacacca atgtggacca ggacctcgtc 2700
ggctggcaag cgccccccgg ggcgcgttcc ttgacaccaa gcacctgcgg cagctcggac 2760
ctttacttgg tcacgaggca tgccgatgtc attccggtgc gccggcgggg cgacagcagg 2820
gggagcctac tctccccag gcccgctctc tacttgaagg gctcttcggg cgtctcactg 2880
ctctgccctt cggggcacgc tgtgggcate tttcgggctg ccgtgtgcac ccgaggggtt 2940
gcgaaggcgg tggactttgt acccgtcgag tetatggaaa ccactatgcy gtccccggtc 3000
ttcacggaca actcgtcccc tccggccgta ccgcagacat tccagggtggc ccatctacac 3060
gcccctactg gtagcggcaa gagcactaag gtgcgggctg cgtatgcagc ccaaggggat 3120
aaggtgcttg tcctgaaccc gtcctgcgcc gccaccctag gtttcgnggc gtatatgtct 3180
aaggcacatg gtatcgaccc taacatcaga accggggtaa ggaccatcac cacgggtgcc 3240
cccatcacgt actccacctc tggcaagttt cttgccgacg gtggttgctc tgggggcggc 3300
tatgacatca taatatgtga tgagtccac tcaactgact cgaccactat cctgggcate 3360
ggcacagctc tggaccaagc ggagcggct ggagcgcgac tcgtcgtgct cgccaccgct 3420
acgcctccgg gatcggtcac cgtgccacat ccaaacatcg aggaggtggc tctgtccagc 3480
actggagaaa tcccccttta tggcaaaagg atccccatcg agaccatcaa gggggggagg 3540

```

55

	cacctcattt	tctgccattc	caagaagaaa	tgtgatgagc	tcgccgcgaa	gctgtccggc	3600
	ctcggactca	atgtctgtagc	atattaccgg	ggccttgatg	tatccgctcat	accaactagc	3660
	ggagacgtca	ttgtcgtagc	aacggacgct	ctaatacgag	gctttaccgg	cgaatttcgac	3720
5	tcagtgtatc	actgcaatac	atgtgtcacc	cagacagtcg	acttcagcct	ggacccgacc	3780
	ttcaccattg	agacgcagac	cgtgccacaa	gacgcggtgt	cacgctcgca	gcggcgaggc	3840
	aggactggta	ggggcaggat	gggcattttac	aggttttgtga	ctccaggaga	acggccctcg	3900
	ggcatgttcg	attcctcggg	tctgtgcgag	tgtatgtacg	cgggctgtgc	ttggtagcag	3960
	ctcacgcccc	ccgagacctc	agttaggttg	cgggcttacc	taaacaacac	aggggtgtccc	4020
	gtctgccagg	accatctgga	gtcttgggag	agcgtcttta	caggcctcac	ccacatagac	4080
	gcccattttc	tgtcccagac	taagcaggca	ggagacaact	tcccctacct	ggtagcatat	4140
10	caggctacgg	tgtgcgccag	ggctcaggct	ccacctccat	cgtgggacca	aatgtggaag	4200
	tgtctcatal	ggctaagacc	tacgtgcac	gggccaacgc	ccctgctgta	taggctggga	4260
	gcccgttcaaa	acgagggttac	taccacacac	cccataacca	aatacatcat	ggcatgcatg	4320
	tcggctgacc	tggagggtcgt	cacgagcacc	tgggtgctgg	taggcggagt	cctagcagct	4380
	ctggcccgct	attgcctgac	aacaggcagc	gtggctcatg	tgggcaggat	catcttgctc	4440
	ggaaagccgg	ccatcattcc	cgacaggga	gtcctttacc	gggagttcga	tgagatggaa	4500
	gagtgccgct	cacacctccc	ttacatcgaa	cagggaatgc	agctcgccga	acaattcaaa	4560
15	cagaaggcaa	tcgggttgct	accaaagcgc	cggaagctgc	cggaagctgc	tgctcccgta	4620
	gtggaaatcca	agtggcgagc	cctcgaagcc	ttctgggcga	agcatagttg	gaatttcatt	4680
	agcgggatac	aataatttagc	aggcttgtcc	actctgcctg	gcaaccccg	gatagcatca	4740
	ctgatggcat	tcacagcctc	tatcaccagc	ccgctcacca	cccaacatac	cctcctgttt	4800
	aacatcctgg	ggggatgggt	ggcgcccaaa	cttgctcttc	ccagegctgc	ttctgctttc	4860
	gtaggcgccg	gcatcgctgg	agcggctgtt	ggcagcatag	gcttgggaa	gggtgctgtg	4920
20	gataattttg	caggttatgg	agcagggttg	gcaggcgccg	tcgtggccct	taaggctcat	4980
	agcggcgaga	tgccctccac	cgaggacctg	gttaacctac	tccctgctat	cctctccctt	5040
	ggcgccctag	tcgtcggggt	cgtgtgcgca	gcgatactgc	gtcggcacgt	ggggccagg	5100
	gagggggctg	tgcatgggat	gaacgggctg	atagcgttcg	cttcgcgggg	taaccacgtc	5160
	tccccacgc	actatgtgcc	tgagagcgac	gctgcagcac	gtgtcactca	gatcctctct	5220
	agtcctacca	tactcagct	gctgaagagg	cttcaccagt	ggatcaacga	ggactgtccc	5280
	acggcatgct	ccggctcgtg	gctaagagat	gtttgggatt	ggatatgcac	gggtgtgact	5340
25	gatttcaaga	cctggctcca	gtccaagctc	ctgccgcgat	tgccgggagt	ccccttcttc	5400
	tcattgtcaac	gtgggtacaa	gggagtctgg	cgggggcgag	gcatcatgca	aaccacctgc	5460
	ccatgtggag	cacagatcac	cgagcatgtg	aaaaacgggt	ccatgaggat	cgtggggcct	5520
	aggacctgta	gtaaacacgt	gcatggaaac	ttccccatta	acgcgtacac	cacgggcccc	5580
	tgcaogccct	ccccggcgcc	aaattattct	agggcgctgt	ggcggttgcc	tgctgaggag	5640
30	tacgtggagg	ttacgcgggt	gggggatttc	cactacgtga	cgggcatgac	cactgacaa	5700
	gtaaagtgcc	cgtgtcaggt	tccggccccc	gaattcttca	cagaagtggg	tgggggtcgg	5760
	ttgcacaggt	acgctccagc	gtgcaaaccc	ctcctacggg	aggaggtcac	attcctggtc	5820
	gggtcctaac	aatacctggt	tgggtcacag	ctcccatgcg	agcccgaacc	ggagctagca	5880
	gtgctcaatt	ccatgctcac	cgacccctcc	cacattacgg	cgagagacgg	taagcgtagg	5940
	ctggccagg	gatctcccc	ctccttgccc	agctcatcag	ctagccagct	gtctgcgctc	6000
	tccttgaagg	caacatgcac	taccctcat	gactccccgg	acgtgacct	catcgaggcc	6060
35	aacctcctgt	ggcggcagg	gatggggggg	aaatcacccc	gcgtggagtc	agaaaaaag	6120
	gtagttaatt	tggaactctt	cgagccgctc	caagcggagg	aggatgagag	ggaagtatcc	6180
	gttccggcgg	agatcctgcg	gaggtccagg	aaattccctc	gagcgtgccc	catatgggca	6240
	cgcccgaggt	acaacccctc	actgttagag	tcctggaagg	acccggacta	cgtccctcca	6300
	gtggtacacg	ggtgtccatt	acggccctgc	aaggccctcc	cgataccacc	tccacggagg	6360
	aagaggacgg	ttgtcctgtc	agaaatctacc	gtgtcttctg	ccttggcgga	gctcgccaca	6420
40	aagaccttcg	gcagctccga	atcgtcggcc	gtcgacagcg	gcacggcaac	ggcctctcct	6480
	gaccagccct	ccgacgacgg	cgcagcgagg	tccgacgttg	agtcgtactc	ctccatgccc	6540
	ccccttgagg	gggagccggg	ggatcccgat	ctcagcgacg	ggctcttggt	taccgtaagc	6600
	gaggaggcta	gtgaggacgt	cgtctgctgc	tcgatgtcct	acacatggac	aggcgccctg	6660
	atcacgcca	gcgctgcgga	ggaaaccaa	ctgcccatac	atgcactgag	caactctttg	6720
	ctccgtcacc	acaacttggt	ctatgctaca	acatctcgca	gcgcaagcct	gcggcagaag	6780
	aaggtcacct	ttgacagact	gcagggtcct	gacgacct	accgggacgt	gctcaaggag	6840
45	atgaagcgga	aggcgtccac	agttaaggct	aaacttctat	ccgtggaggga	agcctgttaa	6900
	ctgacgcccc	cacattcggc	cagatctaaa	tttggtatg	gggcaaagga	cgtccggaac	6960
	ctatccagca	aggcgttaa	ccacatccgc	tccgtgtgga	aggacttgct	ggaagacact	7020
	gagacaccaa	ttgacaccac	catcatggca	aaaaatgagg	tttcttgcgt	ccaaccagag	7080
	aagggggggc	gcaagccagc	tcgccttatc	gtattcccag	atttgggggt	tcgtgtgtgc	7140
	gagaaaaatg	ccctttacga	tgtggtctcc	accctccctc	aggccgtgat	gggctcttca	7200
	tacggattcc	aatactctcc	tggaacagcg	gtcagagttc	tggtgaatgc	ctggaaaagc	7260
50	aagaaatgcc	ctatgggctt	cgcatatgac	acccgctgtt	ttgactcaac	ggtcactgag	7320
	aatgacatcc	gtgttgagga	gtcaatctac	caatgtttgt	acttggcccc	cgaagccaga	7380
	caggccataa	ggtcgctcac	agagcggctt	tacatcgggg	gccccctgac	taattctaaa	7440
	gggcagaact	gcggctatcg	ccggtgccc	gcgagcgggt	tactgacgac	cagctgcggt	7500

EP 1 043 399 A2

5 aataccctca catgttactt gaaggccgct gcggcctgtc gagctgcgaa gctccaggac 7560
 tgacgatgc tcgtatgcgg agacgacctt gtcgttatct gtgaaagcgc ggggacccaa 7620
 gaggacgagg cgagcctacg ggccttcacg gaggctatga ctagatactc tgccccccct 7680
 ggggacccgc ccaaaccaga atacgacttg gagttgataa catcatgtct ctccaatgtg 7740
 tcagtcgcgc acgatgcatc tggcaaaaagg gtgtactatc tcacccgtga cccaccacc 7800
 ccccttgccg gggctgcgtg ggagacagct agacacactc cagtcaattc ctggctaggc 7860
 aacatcatca tgtatgcgcc caccttggtg gcaaggatga tcctgatgac tcatttcttc 7920
 tccatccttc tagctcagga acaacttgaa aaagccctag attgtcagat ctacggggcc 7980
 10 tgttactcca ttgagccact tgacctacct cagatcattc aacgactcca tggccttagc 8040
 gcattttcac tccatagtta ctctccaggc gagatcaata gggtaggttc atgcctcagg 8100
 aaacttgggg taccgccctt gcgagtctgg agacatcggg ccagaagtgt ccgcgctagg 8160
 ctactgtccc agggggggag ggctgccact tgtggcaagt acctcttcaa ctgggcagta 8220
 aggaccaagc tcaaaactcac tccaatcccg gctgcgtccc agttggattt atccagctgg 8280
 15 ttcgttgctg gttacagcgg gggagacata tatcacagcc tgtctcgtgc ccgaccccg 8340
 tggttcatgt ggtgcctact cctactttct gtaggggtag gcattctatct actccccaac 8400
 cgatgaacgg ggagctaaac actccaggcc aataggccat cctgtttttt tccctttttt 8460
 tttttctttt tttttttttt tttttttttt tttttttttt ctcttttttt tttcctcttt 8520
 ttttcctttt ctttcctttg gtggctccat cttagcccta gtcacggcta gctgtgaaag 8580
 20 gtccgtgagc cgcttgactg cagagagtgc tgatactggc ctctctgcag atcaagt 8637
 25
 30
 35
 40
 45
 50
 55

<210> 3
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

5

<400> 3
 gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tcttcacgca gaaagcgctct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
 cccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180
 gacgaccggg tcctttcttg gatcaaccg ctcaatgcct ggagatttgg gcgtgcccc 240
 gcgagactgc tagccgagta gtgttgggtc gcgaaggcc ttgtgggtact gcctgatagg 300
 ctcaaagaaa aaccaaactg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
 acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
 caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 540
 ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
 cgtggctggc caccgacggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
 gaagggactg gctgctattg ggcgaagtgc cggggcaggga tctcctgtca tctcaccctt 720
 ctctctgccg gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
 cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
 tggaaagcgg tcttgtcgat caggatgac tggacgaaga gcatcagggg ctgcgcccag 900
 ccgaactgtt cgccaggctc aaggcgcgca tgcccgacgg caggatctc gtctgaccc 960
 atggcgatgc ctgcttgcgg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
 actgtggcgg gctgggtgtg gccgaccgct atcaggacat agcgttggct acccttgata 1080
 ttgtcgaaga gcttggcggc gaatgggctg accgcttctc cgtgctttac ggtatgcggc 1140
 ctcccgatcc gcagcgcatc gccttctatc gccttcttga cgagtctctc tgagttctaa 1200
 cagaccacaa cgggttccct ctacgggat caattccgcc cctctccctc cccccccct 1260
 aacgttactg gccgaagcgg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
 tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
 acgagcattc ctagggtctc tcccctctc gccaaaggaa tgcaaggctc gttgaatgtc 1440
 gtgaaggaa cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
 tgacggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
 taagataccc ctgcaaaagg ggcacaaccc cagtgccacg ttgtgagttg gatagtgtg 1620
 gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
 gtaccccatc gtatgggatc tgatctgggg cctcggtgca catgctttac atgtgtttag 1740
 tcgaggttaa aaaacgtcta gggcccccga accacgggga cgtgggtttc ctttgaaaaa 1800
 cagcataata ccatggcgcc tattaccggc tactcccaac agacgagagg cctacttggc 1860
 tgcatcatca ctacgctcac aggcggggac aggaaccagg tcgaggggga ggtccaaagt 1920
 gtctccaccg caacacaatc ttctctggcg acctgcgtca atggcggtgt ttggactgtc 1980
 tatcatgggt ccggctcaaa gacccttgcc ggcccaagg gcccaatcac ccaaatgtac 2040
 accaatgtgg accaggacct cgtcggctgg caagcgcccc ccggggcgcg ttccttgaca 2100
 ccattgcacct gcggcagctc ggacctttac ttgttcacga ggcattgccg tgcattccc 2160
 gtgcgcccgc ggggcgacag cagggggagc ctactctccc ccaggcccg ctactacttg 2220
 aagggtctct cgggcgggtc actgctctgc ccctcggggc acgctgtggg catcttctcg 2280
 gctgacgtgt gcaccgagg ggttgcgaag gcggtggact ttgtaccctg cgagtctatg 2340
 gaaaccacta tgccgtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgag 2400
 acattccagg tgcccatct acacgcccct actggtagcg gcaagagcac taagggtgcc 2460
 gctgcgtatg cagcccaagg gtataagggt cttgtcctga acccgctccg cgccgccacc 2520
 ctagggtttc gggcgatatg gtctaaggca catggtatcg accctaaccat cagaaccggg 2580
 gtaaggacca tcaccacggg tgcccccatc acgtactcca cctatggcaa gtttcttgcc 2640
 gacgggtggt gctctggggg cgccatgac atcataatat gtgatgagtg ccactcaact 2700
 gactcgacca ctatcctggg catcggcaca gtctggacc aagcggagac ggctggagcg 2760
 cgactcgtcg tgctcgccac cgctacgctc ccgggatcgg tcaccgtgcc acatccaaac 2820
 atcgaggagg tggctctgtc cagcactgga gaaatccct tttatggcaa agccatcccc 2880
 atcgagacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940
 gagctcgccg cgaagctgtc cggcctcgga ctcaatgctg tagcatatta ccggggcctt 3000
 gatgtatccg tcataccaac tagcggagac gtcattgtcg tagcaacgga cgctctaagt 3060
 acgggcttta ccggcgattt cgactcagtg atcgactgca atacatgtgt caccagaca 3120
 gtgcacttca gcctggacce gaccttcacc attgagacga cgaccgtgcc acaagacgg 3180
 gtgtcacgct cgcagcggcg aggcaggact ggtaggggca ggatgggcat ttacagggtt 3240
 gtactccag gagaacggcc ctccggcatg ttcgattcct cgggtctgtg cgagtctat 3300
 gacgcgggct gtgcttggta cgagctcacg cccgccgaga cctcagttag gttgcgggct 3360
 tacctaaca caccaggggt gccgtctgc caggaccatc tggagtctcg ggagagcgtc 3420
 tttacaggcc tcaccacat agacgcccac ttctgttccc agactaagca ggcaggagac 3480
 aacttcccc acctggtagc ataccaggt acggtgtgcg ccagggtcga ggctccacct 3540

55

EP 1 043 399 A2

5

10

15

20

25

30

35

40

45

50

55

```

ccatcgtggg accaaatgtg gaagtgtctc atacggctaa agcctacgct gcacggggcca 3600
acgccccctgc tgtataggtc gggagccggt caaaacgagg ttactaccac acaccccata 3660
accaaataca tcatggcatg catgtcggct gacctggagg tcgtcacgag cacctgggtg 3720
ctggtagggc gagtcctagc agctctggcc gcgtattgcc tgacaacagg cagcgtggtc 3780
atttgtgggca ggatcatctt gtccggaaag ccggccatca tccccgacag ggaagtcctt 3840
taccgggagt tcgatgagat ggaagagtgc gcctcacacc tcccttacat cgaacaggga 3900
atgcagctcg ccgaacaatt ccaacagaa gcaatcgggt tcttgcaaac agccaccaag 3960
caagcggagg ctgctgctcc cgtggtggaa tccaagtggc ggaccctcga agccttctcg 4020
gcgaagcata tgtggaattt catcagcggg atacaatatt tagcaggctt gtccactctg 4080
cctggcaacc ccgcgatagc atcactgatg gcattcacag cctctatcac cagcccgctc 4140
accacccaac ataccctcct gtttaacatc ctggggggat gggtagccgc ccaacttget 4200
cctccacggc ctgctctcgc tttcgtaggc gccggcatcg ctggagcggc tgttggcagc 4260
ataggccctt ggaaggtgct tgtggatatt ttggcaggtt atggagcagg ggttggcagg 4320
gcgctcgttg cctttaaggt catgagcggc gagatgccct ccaccgagga cctggttaac 4380
ctactccctg ctatccctct cctggcgccc ctagtctcgc gggctcgttg cgcagcgata 4440
ctgctcgtgc acgtgggccc aggggagggg gctgtgcagt ggaatgaacc gctgatagcg 4500
ttcgtctcgc ggggttaacca cgtctccccc acgcactatg tgcctgagag cgacgctgca 4560
gcacgtgtca ctacagatct accatcactc agctgctgaa gaggcttacc 4620
cagtggatca acgaggaact ctcacgcca tgcctcggct cgtggctaag agatgtttgg 4680
gattggatat gcacgggtgtt gactgatttc aagacctggc tccagtccaa gctcctgccc 4740
cgattgcccg gattcccctt cttctcatgt caacgtgggt acaagggaat ctggcggggc 4800
gacggcatca tgcaaaaccac ctgcccattgt ggagcacaga tcaccggaca tgtgaaaaac 4860
ggttccatga ggatcgtggg gcctaggacc tgtagtaaca cgtggcatgg aacattcccc 4920
attaacgcgt acaccacggg cccctgcacg ccctcccccg cgcgaattta tctaggcgcg 4980
ctgtggcggg tggctgctga ggagtacgtg gaggttacgc gggtagggga tttccactac 5040
gtgacgggca tgaccactga caacgtaaag tgcccgtgtc aggttccggc ccccgaaattc 5100
ttcacagaag tggatgggggt gcggttgac aggtacgctc cagcgtgcaa accctccta 5160
cgggaggagg tcacattcctt ggtcgggctc aatcaatacc tggttgggtc acagctccca 5220
tgcgagcccg aaccggagct agcagtgtct acttccatgc tcaccgaccc ctcccacatt 5280
acggcggaga cggctaagcg taggctggcc aggggatctc cccctcctt ggccagctca 5340
tcagctagcc agctgtctgc gccttccctg aaggcaacat gcactaccog tcatgactcc 5400
ccggacgctg acctcatcga ggccaacctc ctgtggcggc aggagatggg cgggaacatc 5460
acccgctgg agtcagaaaa taaggtagta attttgact ctttcgagcc gctccaagcg 5520
gaggaggatg agagggaagt atccgttccg gcggagatcc tgcggaggtc caggaaattc 5580
cctcgagcga tgcccatatg ggcacgcccg gattacaacc ctccactgtt agagtcttgg 5640
aaggacctcg actacgtccc tccagtggta caggggtgtc cattgcccgc tgccaaggcc 5700
cctccgatac cacctccacg gaggaagagg acggttgtcc tgcagaatc taccgtgtct 5760
tctgccttgg cggagctcgc cacaagacc ttcggcagct ccgaatcgtc ggccgtcgac 5820
agcgcgacgg caacggcctc tctgaccag ccctccgacg acggcgacgc gggatccgac 5880
ggttagtctg actcctccat gccccccctt gagggggagc cgggggatcc cgatctcagc 5940
gacgggtctt ggtctaccgt aagcgaggag gctagttagg acgtcgtctg ctgctcgtatg 6000
tcctacacat ggacaggcgc cctgatcacg ccatgcgctg cggaggaaac caagtctccc 6060
atcaatgcac tgagcaactc tttgtccctt caccacaact tgggtctatgc tacaactctc 6120
cgacgcgcaa gcctgcggca gaagaaggct acctttgaca gactgcaggc cctggacgac 6180
cactaccggg acgtgctcaa ggagatgaag gcgaaggcgt ccacagttaa ggctaaactt 6240
ctatccgtgg aggaagcctg taagctgacg ccccacatt cggccagatc taaatttggc 6300
tatggggcaa aggaagtcct gaacctatcc agcaaggccg ttaaccacat ccgctccgtg 6360
tggaaggact tgctggaaga cactgagaca ccaattgaca ccaccatcat ggcaaaaaat 6420
gaggttttct gcgtccaacc agagaagggg ggccgcaagc cagctcgcct tatcgtattc 6480
ccagatttgg gggttcgtgt gtgcgagaaa atggcccttt acgatgtggg ctccaccctc 6540
cctcaggccg tgatgggctc ttcatacgga ttccaatact ctccctggaca gcgggtcgag 6600
ttcctggtga atgcctggaa agcgaagaaa tgccctatgg gcttcgcata tgacaccgcg 6660
tgttttgact caacgggtcac tgagaatgac atccgtgttg aggagtcaat ctaccaatgt 6720
tgtgacttgg ccccggaagc cagacaggcc ataaggctgc tcacagagcg gctttacatc 6780
gggggccccg tgactaattc cggtaatacc aactgcggct atcgccgtg ccgcgcgagc 6840
ggtgtactga cgaccagctg cgtcaatacc ctcacatgtt acttgaaggc cgtgcggccc 6900
tgtcagctg cgaagctcca ggactgcacg atgtcgtat gcggagacga ccttgtcgtt 6960
atctgtgaaa gcgcggggac ccaagaggac gaggcgagcc tacgggcctt cacggaggct 7020
atgactagat actctgcccc ccttggggac cgcaccaaac cagaatacga cttggagtgt 7080
ataacatcat gctcctccaa tgtgtcagtc gcgcacgat catctggcaa aagggtgtac 7140
tatctcaccg gtgaccccac caccctctt gcgcgggctg cgtgggagac agctagacac 7200
actccagta attcctggct agccaacatc atcatgtatg cgccacctt gtgggcaagg 7260
atgactctga tgactcattt cttctccatc cttctagctc aggaacaact tgaaaaagcc 7320
ctagattgtc agatctacgg ggctctgtac tccattgagc cacttgacct acctcagatc 7380
attcaacgac tccatggcct tagcgcattt tcactccata gttactctcc aggtgagatc 7440
aatagggtgg cttcatgcct caggaaactt ggggtaccgc ccttgcgagt ctggagacat 7500

```

EP 1 043 399 A2

	cgggccagaa	gtgtccgcgc	taggtactg	tcccaggggg	ggagggctgc	cacttggtgc	7560
	aagtacctct	tcaactgggc	agtaaggacc	aagctcaaac	tcactccaat	cccggtgcg	7620
	tcccagttgg	atztatccag	ctggttcgtt	gctggttaca	gcgggggaga	catatatcac	7680
5	agcctgtctc	gtgcccgacc	ccgctgggtc	atgtggtgcc	tactcctact	ttctgtaggg	7740
	gtaggcacat	atctactccc	caacegatga	acggggagct	aaacactcca	ggccaatagg	7800
	ccatcctggt	tttttccctt	tttttttttc	tttttttttt	tttttttttt	tttttttttt	7860
	ttttctcctt	tttttttctt	ctttttttcc	ttttctttcc	tttggtggct	ccatcttagc	7920
	cctagtcacg	gctagctgtg	aaaggtccgt	gagccgcttg	actgcagaga	gtgctgatac	7980
10	tggcctctct	gcagatcaag	t				8001

15

20

25

30

35

40

45

50

55

<210> 4
 <211> 7989
 <212> DNA
 <213> Hepatitis C Virus

5

```

<400> 4
gccagccccc gattgggggg gacactccac catagatcac tccccgtgta ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggctcg cggaaaccgt gagtacaccg gaattgccag 180
gacgaccggg tctttcttg gatcaaccg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtaact gcctgatagg 300
gtgcttgca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
tcaaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc 420
cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct 480
ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgcc gggtctttt gtcaagaccg 540
acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca 600
cgacggcggt tccttgccga gctgtgctcg acgttgtcac tgaagcggga agggactggc 660
tgctattggg cgaagtggcg gggcaggatc tctgtcatc tcacctgct cctgccaga 720
aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc 780
cattcgacca ccaagcgaata catcgcatcg agcagcacg tactcggatg gaagccggtc 840
ttgtcgatca ggaatgatcg gacgaagagc atcaggggct cgcgccagcc gaactgttcg 900
ccaggctcaa ggctgcctg cccgacggcg aggatctcgt cgtgaccat ggcatgcct 960
gcttgccgaa tatcatggtg gaaaatggcc gctttcttgg attcatcgac tgtggccggc 1020
tgggtgtggc ggaatgatcg cggcagcgat attcggctac cgtgatatt gctgaagac 1080
ttggcggcga atgggctgac cgtctctcgt tgccttaccg tatcgccgt cccgattcgc 1140
agcgcactcg cttctatcgc cttcttgacc agttctctcg agtttaaca gaccacaacc 1200
gtttccctct agcgggatca attccgcccc tctccctccc cccccctaa cgttaactgg 1260
cgaagccgct tggaaataagg ccggtgtgctg tttgtctata tgttattttc caccataattg 1320
ccgtcttttg gcaatgtgag ggcgggaaa cctggccctg tcttcttgac gaggattcct 1380
aggggtcttt cccctctcgc caaaggaatg caaggctgtg tgaatgtcgt gaaggaagca 1440
gttctcttgg aagcttcttg aagacaaaca acgtctgtag cgaccctttg caggcagcgg 1500
aaccctccac ctggcgacag gtgcctctgc ggccaaaagc cactgtgata agatacact 1560
gcaagggcg cacaaccaca gtgcacggt gtgagttgga tagttgtgga aagagtcaca 1620
tggctctcct caagcgtatt caacaagggg ctgaaggatg ccagaaagg accccattgt 1680
atgggatctg atctggggcc tcggtgcaca tgcctttacat gtgttttagt gagggttaaa 1740
aacgtctagg cccccgaac caggggacg tggttttcct ttgaaaaaca cgataaact 1800
atggcgccca ttacggccta ctcccaacag acgcgaggcc tacttggtg catcatcact 1860
agcctcacag gcccgggacag gaaccaggtc gaggggagg tccaagtggt ctccaccgca 1920
acacaatctt tcctggcgac ctgcgtcaat ggcgtgtgtt ggactgtcta tcattgtgccc 1980
ggctcaaaaga ccttgccgg cccaaagggc ccaatcacc aaatgtacac caatgtggac 2040
caggacctcg tcggttgcca agcggcccc ggggcgctt ccttgacacc atgcacctgc 2100
ggcagctcgg acctttactt ggtcacgagg catgcgatg tcattccggt gcgcccggc 2160
ggcgacagca gggggagcct actctcccc agggccgtct cctacttgaa gggtcttcg 2220
ggcggtccac tgcctgccc ctccggggcac gctgtgggca tctttcgggc tgcgtgtgc 2280
acccgagggg ttgcgaaggc ggtggacttt gtaccgtcg agtctatgga aaccactatg 2340
cggctccccg tcttcacgga caactcgtcc cctccggcgg taccgagac attccaggtg 2400
gcccattctac acgcccctac tggtagcggc aagagcacta aggtgccggc tgcgtatgca 2460
gccccagggt ataagggtgt tgcctgaac ccgtccgtcg ccgccaccct aggtttcggg 2520
gcgtatatgt ctaaggcaca tggtagcag cctaaccatc gaaccgggt aaggaccatc 2580
accacgggtg ccccatcac gtactccacc tatggcaagt ttcttgccga cgggtggtgc 2640
tctggggggc cctatgacat cataatatgt gatgagtgcc actcaactga ctgaccact 2700
atcctgggca tcggcacagt cctggaccaa gcgagacgg ctggagcgcg actcgtcgtg 2760
ctcgccaccg ctacgcctcc gggatcgtgc accgtgccac atccaaaacat cgaggagggtg 2820
gctctgtcca gcaactggaga aatccccctt tatggcaaa aatgtgatga gctcgccggc 2880
aaggggggga ggcacctcat tttctgcat tccaagaaga aatgtgatga gctcgccggc 2940
aagctgtccg gcctcggact caatgctgta gcatattacc ggggccttga tgtatccgtc 3000
ataccaacta gcggagacgt cattgtcgta gcaacggacg ctctaatac gggctttacc 3060
ggcgatttcg actcagtgat cgactgcaat acatgtgtca ccagacagt cgacttcagc 3120
ctggacccga ccttcacatc taggagcagc accgtgccac aagacgcggg gtacgctcg 3180
cagcgccgag gcaggactgg taggggcagg atgggcattt acaggtttgt gactccagga 3240
gaacggccct cgggcatgtt cgattcctcg gttctgtcgg agtgctatga cgcgggctgt 3300
gcttggtacg agctcacgcc cgccgagacc tcagttaggt tgcgggctta cctaaacaca 3360
ccagggttgc ccgtctgcca gaccatctg gagttctggg agagcgtctt tacaggcctc 3420
acccacatag acgcccattt cttgtccag actaagcagg caggagacaa ctccctac 3480
ctggtagcat accaggctac ggtgtgccc agggctcagg ctccacctcc atcgtgggac 3540

```

55

EP 1 043 399 A2

5
10
15
20
25
30
35
40
45
50
55

caaatgtgga agtgtctcat acggctaaga cctacgctgc acggggccaac gccctgtctg 3600
tataggctgg gagccgttca aaacgaggtt actaccacac acccataaac caaatatcatc 3660
atggcatgca tgcggctga cctggaggtc gtcacgagca cctgggtgct ggtaggcgga 3720
gtccttagcag ctctggccgc gtattgcctg acaacaggca gcgtggctcat tgtgggcagg 3780
atcatcttgt ccggaaagcc ggccatcatt cccgacaggg aagtccctta ccgggagttc 3840
gatgagatgg aagagtgcgc ctccacacct ccttacatcg aacagggaat gcagctcgcc 3900
gaacaattca aacagaaggc aatcgggttg ctgcaaacag ccaccaagca agcggaggct 3960
gctgctcccc tgggtggaatc caagtggcgg accctcgaag ccttctgggc gaagcatatg 4020
tggaaattca tcagcgggat acaatattta gcaggcttgt ccactctgcc tggcaacccc 4080
gcgatagcat cactgatggc attcacagcc tctatcacca gcccgctcac caccacat 4140
accctctctg ttaacatcct ggggggtagg gtggccgccc aacttgctcc tcccagcgt 4200
gcttctgctt tcgtaggcgc cgccatcgct ggagcggctg ttggcagcat aggccttggg 4260
aagggtgctt tggatatttt ggcaggttat ggagcagggg tggcagggcg gctcgtggcc 4320
tttaaggtca tgagcggcga gatgccctcc accgaggacc tgggttaacct actccctgct 4380
atcctctccc ctggcgcccc agtcgtcggg gtcgtgtgcg cagcgatact gcgtcggaac 4440
gtgggcccag gggagggggc tgtgcagtgg atgaaccggc tgatagcgtt cgcttcggcg 4500
ggtaaccacg tctccccac gcactatgtg cctgagagcg acgtgcgagc acgtgtcact 4560
cagatctctc ctagtcttac ctgactcag ctgctgaaga ggcttcacca gtggatcaac 4620
gaggactgct ccacgccatg ctccggctcg tggctaagag atgtttggga ttggatatgc 4680
acggtgttga ctgatttcaa gacctggctc cagtccaagc tctgcggcg atttgcggga 4740
gtccccctct tctcatgtca acgtgggtac aaggaggtct ggccggggcg cggtcatcatg 4800
caaacaccct gcccatgtgg agcacagatc accggacatg tgaaaaacgg tcccatgagg 4860
atcgtggggc ctaggacctg tagtaaacag tggcatggaa cattccccat taacgcgtac 4920
accacgggcc cctgcacgccc ctcccggcg ccaaatattt ctaggggcgt gtggcgggtg 4980
gctgctgagg agtacgtgga ggttacgccc gtgggggatt tccactacgt gacgggcatg 5040
accactgaca acgtaaagtg cccgtgtcag gttccggccc ccgaattctt cacagaagtg 5100
gatgggggtg ggttgacag gtacgctcca gcgtgcaaac ccctcctacg ggaggaggtc 5160
acattcctgg tcgggctcaa tcaatacctg gttgggtcac agtcccatg cgagcccgaa 5220
ccggacgtag cagtgtcac ttccatgctc accgacccct ccacattac ggcggagacg 5280
gctaagcgta ggtgtggcag gggatctccc cctccttgg ccagctcatc agtccggagc 5340
ctgtctgccc ctctcctgaa ggcaacatgc actaccgctc atgactcccc ggacgtgac 5400
ctcatcgagg ccaacctcct gtggcgccag gagatggcg ggaacatcac ccgctggag 5460
tcagaaaaata aggtagtaat tttggactct ttccagccgc tccaagcgga ggaggatgag 5520
agggaagtat ccgttccggc ggagatcctg cggagggtcca ggaattcccc tcgagcgatg 5580
cccatatggg cagcccgga ccaactgttag ccactgttag agtccctgga ggaccggac 5640
tacgtccctc cagtgttaca cgggtgtcca ttgcccctg ccaaggcccc tccgatacca 5700
cctccacgga ggaagaggac ggttgtcctg tcagaaatcta ccgtgtcttc tgccttggcg 5760
gagctcccca caaagacctt cgcagctcc gaatcgtcgg ccgtcgacag cggcagggac 5820
acggcctctc ctgaccagcc ctccgacgac ggcgacgccc gatccgacgt tgagtcgtac 5880
tctctcatgc cccccctga gggggagccc ggggatcccc atctcagcga cgggtcttgg 5940
tctacccgta gcgaggagcc tagtgaggac gtcgtctgct gctcgatgtc ctacacatgg 6000
acaggcgccc tgatcacgcc atgcccgtcg gaggaacca agctgccccat caatgcaactg 6060
agcaactctt tgcctccgta ccacaacttg gtctatgcta caacatctcg cagcgcaagc 6120
ctgcggcaga agaaggtcac ctttgacaga ctgcaggtcc tggacgacca ctaccgggac 6180
gtgctcaagg agatgaaggc gaaggcgtcc acagttaagg ctaaaactct atccgtggag 6240
gaagcctgta agctgacgcc cccacattcg gccagatcta aatttggtga tggggcaag 6300
gacgtccgga acctatccag caaggccgtt aaccacatcc gctccgtgtg gaaggacttg 6360
ctggaagaca ctgagacacc aattgacacc accatcatgg caaaaaatga ggttttctgc 6420
gtccaaccag agaagggggg ccgcaagcca gctcgcctta tcgtattccc agatttgggg 6480
gttcgtgtgt gcgagaaaaa ggccctttac gatgtggtct ccaccctccc tcaggccgtg 6540
atgggctctt catacggatt ccaatactct cctggacagc gggctcgagtt cctggtgaat 6600
gcctggaaag cgaagaaatg ccctatgggc ttccgatatg acaccgctg ttttgaactc 6660
acgggtcactg agaatgacat ccgtgttgag gagtcaatct accaatgttg tgacttggcc 6720
cccgaagcca gacaggccat aaggctcgtc acagagcggc tttacatcgg gggccccctg 6780
actaatctta aagggcagaa ctgcccgtat cgccggtgcc gcgagagcgg tgtactgacg 6840
accagctcgc gtaataccct cacatgttac ttgaaggccc ctgcccctg tcgagctgcg 6900
aagctccagg actgcacgat gctcgtatgc ggagacgacc ttgtcgttat ctgtgaaagc 6960
gcgggggccc aagaggacga ggcgagccta cgggccttca cggaggctat gactagatac 7020
tctgcccccc ctgggggacc gcccaacca gaatacgact tggagtgtat aacatcatgc 7080
tcttccaatg tgtcagtcgc gcacgatgca tctggcaaaa ggggtgtaact tctcaccctg 7140
gacccccaca ccccccttgc gcgggctgcg tgggagacag ctagacacac tccagtcaat 7200
tcttgcttag gcaacatcat catgtatgcg cccaccttgt gggcaaggat gatcctgatg 7260
actcatttct tctccatcct tctagctcag gaacaacttg aaaaagccct agattgtcag 7320
atctacgggg cctgttactc cattgagcca cttagacctac ctcagatcat tcaacgactc 7380
catggcccta gcgcatttcc actccatagt tactctccag gtgagatcaa taggggtggc 7440
tcatgcctca ggaaacttgg ggtaccgccc ttgcgagctc ggagacatcg ggcagaagt 7500

EP 1 043 399 A2

	gtccgcgcta	ggctactgtc	ccaggggggg	agggctgcc	cttgtggcaa	gtacctcttc	7560
	aactgggcag	taaggaccaa	gctcaaactc	actccaatcc	cggctgcgtc	ccagttggat	7620
5	ttatccagct	ggttcgttgc	tggttacagc	gggggagaca	tatatcacag	cctgtctcgt	7680
	gcccgaaccc	gctggttcat	gtggtgccta	ctcctacttt	ctgtaggggt	aggcatctat	7740
	ctactcccca	accgatgaac	ggggagctaa	acactccagg	ccaataggcc	atcctgtttt	7800
	tttccctttt	tttttttctt	tttttttttt	tttttttttt	tttttttttt	ttctcctttt	7860
	tttttctctt	ttttttcctt	ttcttttcctt	tggtaggtcc	atcttagccc	tagtcacggc	7920
	tagctgtgaa	aggtccgtga	gccgcttgac	tgcagagagt	gctgatactg	gcctctctgc	7980
10	agatcaagt						7989

15

20

25

30

35

40

45

50

55

<210> 5
 <211> 8649
 <212> DNA
 <213> Hepatitis C Virus

5

10

15

20

25

30

35

40

45

50

55

```

<400> 5
gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgctc agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggctcg cggaaccggg gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaacgt aacaccaacg ggccgcgcat gattgaacaa gatggattgc 420
acgcagggtc tccggcgctc tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg cgggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtgtgctgg cagcaggggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg gctgctattg ggcgaagtgc cggggcaggga tctcctgtca tctcactctg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacctg ccattcgac caccaagcga cagctcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctccgcccag 900
ccgaactgtt cgccaggctc aaggcgcgca tgcccgacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgttgcgg aatatcatgg tggaaaatgg ccgcttttct ggattacatc 1020
actgtggcgg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgtggaaga gcttggcggc gaatgggctg accgcttctc cgtgctttac ggatgcgccg 1140
ctcccgatgc gcagcgcatc gcctctcttc gccttcttga cgaagtcttc tgagtttcaa 1200
cagaccacaa cggtttccct cttagcggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tccaccatat tgccttcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc cttagggctc tccccctctc gccaaaggaa tgcaaggctc gttgaatgtc 1440
gtgaagggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaacctt 1500
tgaccccatc ggaaccccc cctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc gccacaaccc cagtgcacag ttgtgagttg gatagttgtg 1620
gaaagagtca aatggtcttc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtaccccatc gtatgggac cccgactat tgcctgtgca catgctttac atgtgtttag 1740
tcgaggttaa aaaacgtcta ggccccccga accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccatggaccg ggagatggca gcatcgtycg gaggcgcggt tttcgtaggt 1860
ctgactctct tgacctgtc accgactat aagctgttcc tcgctaggct catatgggtg 1920
ttacaatatt ttatcaccag ggccgaggca cacttgcaag tgtggatccc cccctcaac 1980
gttcgggggg gccgcgatgc cgtcatcctc ctacgtgctc cgtaccacc agagctaatt 2040
tttaccatca ccaaatctt gctcgccata ctcggtccac tcatgggtgt caggctggtc 2100
ataaccaaag tgccgtactt cgtgcgcga cacgggtcca ttctgtcatg catgctggtg 2160
cggaagggtt ctgggggtca ttatgtccaa atggctctca tgaagtggc cgactgaca 2220
ggtacgtacg tttatgacca tctaccccc ctgcgggact gggcccaacg gggcctacga 2280
gaccttgcgg tggcagttga gcccgctgct ttctctgata tggagaccaa ggttatcacc 2340
tgggggggcag acaccggcgc gtgtggggac atcatcttgg gctgcccctg ctccgcccg 2400
aggggggagg agatacatct gggaccggca gacagccttg aagggcaggg gtggcgactc 2460
ctcgcccta ttacggccta ctcccaacag acgcgaggcc tacttggctg catcatcact 2520
agcctcacag gccgggacag gaaccaggtc gagggggagg tccaagtgtt ctccaccgca 2580
acacaatctt tcctggcgac ctgcgtcaat ggctgtgtgt ggactgtcta tcatggtgcc 2640
ggctcaaaga ccttgcggc cccaaagggc ccaatcacc aaatgtacac caatgtggac 2700
caggacctcg tcggctggca agcgcccccc ggggcgcgtt ccttgacacc atgcacctgc 2760
ggcagctcgg accttactt ggtcacgagg catgccgatg tcattccggt gcgccggcg 2820
ggcgacagca gggggagcct actctcccc gggccgctct cctaactgaa gggctcttcg 2880
ggcggtccac tgctctgccc ctcggggcac gctgtgggca tcttctgggc tgccgtgtgc 2940
acccgagggg ttgcgaaggc ggtggacttt gtacccgtcg agtctatgga aacctatg 3000
cggctccccg tcttcacgga caactcgtcc cctccggccg taccgcagac attccagggt 3060
gcccactac acgcccctac tggtagcggc aagagcacta aggtgccggc tgcgtatgca 3120
gcccagggt ataaggtgct tgtcctgaac ccgtccgtcg ccgccaccct aggtttcggg 3180
gcgtatatgt ctaaggcaca tggatcgac cctaacatca gaaccgggt aaggaccatc 3240
accaagggt ccccatcac gtactccacc tatggcaagt ttcttgccga cgggtggtgc 3300
tctggggcg cctatgacat cataatatgt gatgagtgc actcaactga ctgaccact 3360
atcctgggca tcggcacagt cctggaccaa gcggagacgg ctggagcgcg actcgtctg 3420
ctcgccaccg ctacgcctcc gggatcggtc accgtgccac atccaaacat cgaggagggt 3480
gctctgtcca gcactggaga aatccccctt tatggcaaa ccatccccat cqagaccatc 3540

```

	aaggggggga	ggcacctcat	tttctgccat	tccaagaaya	aatgtgatya	gctcgccgcg	3600
	aagctgtccg	gcctcggact	caatgctgta	gcataattacc	ggggccttga	tgtatccgtc	3660
	ataccaacta	gcggagacgt	cattgtcgta	gcaacggacg	ctctaatagac	gggctttacc	3720
5	ggcgatttcg	actcagtgat	cgactgcaat	acatgtgtca	cccagacagt	cgacttcagc	3780
	ctggacccca	ccttcacccat	tgagacgacg	accgtgccac	aagacgcggg	gtcacgctcg	3840
	cagcggcgag	gcaggactgg	taggggcagg	atgggcattt	acaggtttgt	gactccagga	3900
	gaacggccct	cgggcatgtt	cgattcctcg	gttctgtgcg	agtgcataga	cgcggtgtgt	3960
	gcttggtacg	agctcacgcc	cgccgagacc	tcagttaggt	tgcggttcta	cctaaacaca	4020
	ccagggttgc	ccgtctgccca	ggaccatctg	gagttctggg	agagcgtctt	tacaggcctc	4080
	acccacatag	acgcccattt	cttgtccccag	actaagcagg	caggagacaa	cttcccctac	4140
10	ctggtagcat	accaggctac	gggtgtgcgc	agggctcagg	ctccacctcc	atcgtgggac	4200
	caaagtgtga	agtgtctcat	acggctaaag	cctacgtctg	acggggccaac	gcccctgctg	4260
	tataggctgg	gagccgttca	aaacgaggtt	actaccacac	accccataac	caaatatcat	4320
	atggcatatg	tgtcggctga	cctggaggtc	gtcacgagca	cctgggtgct	ggtaggcgga	4380
	gtcctagcag	ctctggccgc	gtattgcctg	acaacaggca	gcgtgggtcat	tgtgggcagg	4440
	atcatcttgt	ccggaaagcc	ggccatcatt	cccagacagg	aagtccttta	ccgggagttc	4500
15	gatagatgag	aagagtgcgc	ccttacatcg	aacagggaat	gcagctcgcc	gagctcgcc	4560
	gaacaattca	aacagaaggg	aatcgggttg	ctgcaaacag	ccaccaagca	agcggaggct	4620
	gctgtctccc	tggttggaatc	caagtggcgg	accctcgaag	ccttctgggc	gaagcatatg	4680
	tggaatttca	tcagcgggat	acaatattta	gcaggcttgt	ccaactcgcc	tggcaacccc	4740
	gcgatagcat	cactgatggc	attcacagcc	tctatcacca	gcccgtcac	caccaacat	4800
	acccctcctgt	ttaacatcct	ggggggatgg	gtggccgccc	aacttgctcc	tcccagcgct	4860
20	gcttctgtct	tcgtaggcgc	cgccatcgct	ggagcggctg	ttggcagcat	aggccttggg	4920
	aaggtgcttg	tggatatattt	ggcaggttat	ggagcagggg	tggcagggcg	gctcgtggcc	4980
	tttaagggtca	tgagcggcga	gatgcctccc	accgaggacc	tggttaacct	actccctgct	5040
	atcctctccc	ctggcgccct	agtctgcggg	gtcgtgtcgc	cagcgatact	gcgtcggcac	5100
	gtgggcccag	gggagggggc	tgtgcagtg	atgaaccggc	tgatagcgtt	cgcttcgcgg	5160
	ggtaaccacg	tctcccccac	gcactatgtg	cctgagagcg	acgctgcagc	acgtgtcact	5220
	cagatcctct	ctagtcttac	catcactcag	ctgctgaaga	ggcttcacca	gtggatcaac	5280
25	gaggactgct	ccacgcccag	ctccggctcg	tggctaagag	atgtttggga	ttggatatgc	5340
	acggtgttga	ctgatttcaa	gacctgggtc	cagtccaaag	tccgtgcggc	attggccggga	5400
	gtccccttct	tctcatgtca	acgtgggtac	aaggggagtct	ggcggggcga	cgccatcatg	5460
	caaaccacct	gcccattgtg	agcacagatc	accggacatg	tgaaaaaagg	ttccatgagg	5520
	atcgtggggc	ctaggacctg	tagtaaacag	tggcatggaa	cattccccat	taacgcgtac	5580
	accacggggc	cctgcacgcc	ctcccggcg	ccaaattatt	ctagggcgct	gtggcggtg	5640
30	gctgtctgag	agtaacgtga	ggttacgcgg	gtgggggatt	tccaactacgt	gacgggcatg	5700
	accactgaca	acgtaaagt	cccgtgtcag	gttccggccc	ccgaattctt	cacagaagt	5760
	gatgggggtg	gggtgcacag	gtacgtctca	gcgtgcaaac	ccttcctacg	ggaggaggtc	5820
	acattccttg	tcgggctcaa	tcaatacctg	gttgggtcac	agctcccatg	cgagcccgaa	5880
	ccggacgtag	cagtgtctac	ttccatgctc	accgacccct	cccacattac	ggcggagacg	5940
	gctaagcgta	gggtggccag	gggatctccc	ccctccttgg	ccagctcacc	agctagccag	6000
	ctgtctgcgc	cttcccttga	ggcaacatgc	actacccgtc	atgactcccc	ggacgctgac	6060
35	ctcatcagag	ccaacctcct	gtggcgccag	gagatggggc	ggaacatcac	ccgctgggag	6120
	tcagaaaaata	aggtagtaat	ttcgagcctc	ttcgagccgc	tccaagcgga	ggaggatgag	6180
	agggaagtat	ccgttccggc	ggagatcctg	cgaggttcca	ggaaattccc	tcgagcgatg	6240
	cccatatggg	cacgcccggg	ttacaacctc	ccactgttag	agtcctggaa	ggaacccggg	6300
	tacgtccctc	cagtgggtaca	cggtgtgtcca	ttgccgcctg	ccaaggcccc	tccgatacca	6360
	cctccacgga	ggaaaggagc	gggtgtcctg	tcagaatcta	ccgtgtcttc	tgccttggcg	6420
40	gagctcgcca	caaagacctt	cggcagctcc	gaatcgtcgg	ccgtcgacag	cggcacggga	6480
	acggcctctc	ctgaccagcc	ctccgacgac	ggcgacggcg	gatccgacgt	tgagtcgtac	6540
	tcctccatgc	cccccttga	gggggagccg	ggggatcccg	atctcagcga	cgggtcttgg	6600
	tctaccgtaa	gcgaggaggg	tagtgaggac	gtcgtctgct	gctcgatgtc	ctacacatgg	6660
	acaggcgccc	tgatcacgcc	atgcgctgcg	gaggaaccca	agctgccccat	caatgcaactg	6720
	agcaactctt	tgctccgtca	ccacaacttg	gtctatgcta	caacatctcg	cagcgcaagc	6780
45	ctgcggcaga	agaagggtcac	ctttgacaga	ctgcaggtcc	tggacgacca	ctaccgggac	6840
	gtgctcaagg	agatgaaggc	gaaggcgctc	acagttaagg	ctaaacttct	atccgtggag	6900
	gaagcctyga	agctgacgcc	cccacattcg	gccagatcta	aatttggcta	tggggcaaa	6960
	gacgtccgga	acctatccag	caaggccgtt	aaccacatcc	gctccgtgtg	gaaggacttg	7020
	ctggaagaca	ctgagacacc	aattgacacc	accatcatgg	caaaaaatga	ggttttctgc	7080
	gtccaaccag	agaagggggg	ccgcaagcca	gctcgctcta	tcgtattccc	agatttgggg	7140
	gttcgtgtgt	gcgagaaaat	ggccctttac	gatgtggtct	ccaccctccc	tcaggccgtg	7200
50	atgggctctt	catacggatt	ccaatactct	cctggacagc	gggtcgagtt	cctggtgaat	7260
	gcctggaaa	cgaagaaaatg	ccctatgggc	ttcgcatatg	acacccgctg	ttttgactca	7320
	acggctcactg	agaatgacat	ccgtgttgag	gagtcfaatct	accaatgttg	tgacttggcc	7380
	cccgaagcca	gacaggccat	aaggtcgctc	acagagcggc	tttacatcgg	gggccccctg	7440
	actaatctta	aagggcagaa	ctgcggctat	cgccggtgcc	gcgcgagcgg	tgtactgacg	7500

EP 1 043 399 A2

	accagctgcg	gtaataccct	cacatgttac	ttgaaggccg	ctgcggcctg	tcgagctgcg	7560
	aagctccagg	actgcacgat	gctcgtatgc	ggagacgacc	ttgtcgttat	ctgtgaaagc	7620
	gcgggggacc	aagaggacga	ggcgagccta	cgggccttca	cggaggctat	gactagatac	7680
5	tctgcccccc	ctgggggacc	gccc aaacca	gaatacgact	tggagttgat	aacatcatgc	7740
	tctccaatg	tgtcagtcgc	gcacgatgca	tctggcaaaa	gggtgtacta	tctcaccctg	7800
	gacccccacca	ccccccctgc	gcgggctgcg	tgggagacag	ctagacacac	tccagtcaat	7860
	tcttggttag	gcaacatcat	catgtatgcg	cccaccttgt	gggcaaggat	gatcctgatg	7920
	actcatttct	tctccatcct	tctagctcag	gaacaacttg	aaaaagccct	agattgtcag	7980
10	atctacgggg	cctgttactc	cattgagcca	cttgacctac	ctcagatcat	tcaacgactc	8040
	catggcctta	gcgcattttc	actccatagt	tactctccag	gtgagatcaa	taggggtggct	8100
	tcatgcctca	ggaaacttgg	ggtaccgccc	ttgcgagtct	ggagacatcg	ggccagaagt	8160
	gtccgcgcta	ggctactgtc	ccaggggggg	agggctgccca	cttgtggcaa	gtacctcttc	8220
	aactgggcag	taaggaccaa	gctcaaatc	actccaatcc	cggctgcgtc	ccagttggat	8280
	ttatccagct	ggttcgttgc	tggttacagc	gggggagaca	tatatcacag	cctgtctcgt	8340
15	gcccgaaccc	gctggttcat	gtggtgccta	ctcctacttt	ctgtaggggt	aggcatctat	8400
	ctactcccca	accgatgaac	ggggagctaa	acactccagg	ccaataggcc	atcctgtttt	8460
	tttccctttt	tttttttctt	tttttttttt	tttttttttt	tttttttttt	ttctcctttt	8520
	tttttctctt	ttttttcctt	ttctttcctt	tggtggctcc	atcttagccc	tagtcacggc	8580
	tagctgtgaa	aggtccgtga	gccgcttgac	tgcagagagt	gctgatactg	gcctctctgc	8640
20	agatcaagt						8649
25							
30							
35							
40							
45							
50							
55							

<210> 6
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

<400> 6
 gccagcccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tcttcacgca gaaagcgctc agccatggcg ttagtatgag tgctgtgcag cctccaggac 120
 ccccccctccc gggagagcca tagtggtctg cggaaaccgt gagtacaccg gaattggccag 180
 gacgaccggg tccctttcttg gatcaaccg ctcaatgcct ggagatttgg gcgtgcccc 240
 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgggtact gcctgatagg 300
 gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
 ctcaaagaaa aaccaaacgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
 acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
 caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt 540
 ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
 cgtggctggc cagcaggggc gttccttgcg cagctgtgct cgacgttgc actgaagcgg 660
 gaaggggactg gctgctattg gcggaagtgc cggggcagga tctcctgtca tctcaccttg 720
 ctctgcccga gaaagtatcc atcatggctg atgcaatgcg gcggtgcat acgcttgatc 780
 cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
 tggaaagccgg tctgtcgat caggatgac tggacgaaga gcacagggg ctccgcgcag 900
 ccgaactgtt cgcagggtc aaggcgcgca tgcccagcgg cagggatctc gtcgtgacct 960
 atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
 actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccggtga 1080
 ttgctgaaga gcttggcgcc gaattgggtg accgcttctc cgtgctttac ggtatcgccg 1140
 ctcccgatc gcagcgcatc gccttctatc gccttcttga cgagttcttc tgagttttaa 1200
 cagaccacaa cgttttccct ctacgggat caattccgcc cctctccctc cccccccct 1260
 aacgttactg gccgaagccg cttggaataa ggcgggtgtg cgtttgtcta tatgttattt 1320
 tccaccatat tgcctctttt tggcaatgtg agggcccgga aaacctggccc tgtcttcttg 1380
 acgagcattc cttaggggtct tccccctctc gccaaaggaa tgcaaggctc gttgaatgtc 1440
 gtgaagggaag cagtctctct ggaagcttct tgaagacaaa caacgtctgt agcgacctt 1500
 tgaggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
 taagatacac ctgcaaaagg ggcacaaacc cagtggccag ttgtgagttg gatagttgtg 1620
 gaaagattca aatggctctc ctcaagcgta ttcaacaagg ggtggaagg tgcccagaag 1680
 gtaccccat gtatgggatc tgcctgggg cctcggtgca catgctttac atgtgtttag 1740
 tccaggttaa aaaaagctta ggcctccgca accacgggga cgtgggtttc ctttgaataa 1800
 cagcataata ccatggcgcc tattacggcc tactcccaac agacggcagg cctactttgg 1860
 tgcatcata ctagcctcac aggcgggac aggaaccagg tcgaggggga ggtccaagt 1920
 gtctccaccg caacacaaatc tttcctggcg acctgctgca atggcgtgtg ttggactgtc 1980
 tatcatgtgt ccggctcaaa gaccttggcc ggcccaagg gcccaatcac ccaaatgtac 2040
 accaatgttg accaggacct cgtcggtgg caagcgcccc ccggggcgcg ttccttgaca 2100
 ccatgcacct gcggcagctc ggacctttac ttggtcacga ggcattgccg tgtcattccg 2160
 gtgcgccccg gggcgacag cagggggagc ctactctccc ccaggcccg ctactacttg 2220
 aagggtctt cggcggttcc atgtcttgc cctcggggc atgtgtgtgg catcttctcg 2280
 gctgcccgtg gcacccgagg ggttgcgagg gcggtggact ttgtaccgt cgagtctatg 2340
 gaaaccata tgccgtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgcag 2400
 acattccagg tggcccatct acacgcccc actggtagcg gcaagagcac taagggtccg 2460
 gctgcgtatg cagcccaaagg gtataagggt cttgtcctga acccgctcgt cgccgccacc 2520
 ctagggtttc gggcgatat gtctaaggca catggtatcg accctaact cagaaccggg 2580
 gtagggaaca tcaccacggg tgccccatc acgtactcca cctatggcaa gtttctgcc 2640
 gacggtggtt gctctggggg cgcctatgac atcataatat gtgatgagtg ccactcaact 2700
 gactcgacca ctatcctggg catcggcaca gtcctggacc aagcggagac ggctggagcg 2760
 cgactcgtcg tgctcgccac cgctacgct ccgggatcgg tcaccgtgcc acatccaaa 2820
 atcgaggagg tggctctgtc cagcactgga gaaatccct tttatggcaa agccatcccc 2880
 atcgcgacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940
 gagctggccg cgaagctatc cggcctcgga ctcaatgctg tagcatatta ccggggcctt 3000
 gatgtatccg tcataccaac tagcggagac gtcattgtcg tagcaacgga cgtctaatg 3060
 acgggcttta ccggcgattt cgactcagtg atcgactgca atacatgtgt caccagaca 3120
 gtgcacttca gcctggacct gaccttcacc attgagacga cgaccgtgcc acaagacgcg 3180
 gtgtcacgct cgcagcgcg aggcaggact ggtaggggca ggatgggcat ttacaggttt 3240
 gtgactccag gagaacggcc ctccggcatg ttcgattcct cggttctgtg cgagtgtat 3300
 gacgcgggct gtgcttggtg cgaagtcacg cccgcgaga cctcagttag gttgcgggct 3360
 tacctaaca caccagggtc gccgtctgc caggaccatc tggagttctg ggagagcgtc 3420
 ttacagggc tcaccacat agacgccat ttctgttccc agactaggca ggcaggagac 3480
 aacttccct acctggtagc ataccaggct acgggtgtgc ccagggctca ggctccacct 3540

	ccatcgtggg	accaaagtgt	ggagtgtctc	atacggctaa	agcctacgct	gcacggggcca	3600
	acgcccctgc	tgtataggct	gggagccgtt	caaaacgagg	ttactaccac	acaccccata	3660
	accaaataca	tcatggcatg	catgtcggct	gacctggagg	tcgtcacgag	cacctgggtg	3720
5	ctggtaggcg	gagtcctagc	agctctggcc	gcgtattgcc	tgacaacagg	cagcgtggct	3780
	attgtgggca	ggatcatctt	gtccggaaag	ccggccatca	ttcccgcagc	ggaagtccct	3840
	taccgggagt	tcgatgagat	ggaagagtgc	gcctcacacc	tcctttacat	cgaacagggg	3900
	atgcagctcg	ccgaacaatt	caaacagaa	gcaatcgggt	tgctgcaaac	agccaccaag	3960
	caagcggagg	ctgctgtccc	cgtgggtgaa	tccaagtggc	ggaccctcga	agccttcttg	4020
	gcgaagcata	tgtggaattt	catcagcggg	atacaatatt	tagcaggcct	gtccactctg	4080
	cctggcaacc	ccgcgatagc	atcactgatg	gcattcacag	cctctatcac	cagcccgctc	4140
10	accacccaac	ataccctcct	gtttaacatc	ctggggggat	gggtggccgc	ccaacttgct	4200
	cctcccagcg	ctgcttctgc	tttcgtaggc	gccggcatcg	ctggagcggc	tggtggcagc	4260
	atagggcctt	ggaaggtgct	tgtggatatt	ttggcagggt	atggagcagg	ggtggcaggc	4320
	gcgctcgtgg	cctttaagg	catgagcggc	gagatgccct	ccaccgagga	cctggttaac	4380
	ctactccctg	ctatcctctc	ccctggcgcc	ctagtctctg	gggtcgtgtg	cgcagcgata	4440
	ctgcgtcggc	acgtggggcc	aggggagggg	gctgtgcagt	ggatgaaccg	gctgatagcg	4500
15	ttcgcttcgc	gggtaacca	cgctccccc	acgcactatg	tgtctgagag	cgacgctgca	4560
	gcacgtgtca	ctcagatcct	ctctagtctt	accatcactc	agctgctgaa	gaggctctac	4620
	cagtggatca	acgaggactg	ctccacgcca	tgctccggct	cgtggctaag	agatgtttgg	4680
	gattggatat	gcacgggtgt	gactgatctt	aagacctggc	tcagctccaa	gctcctgccc	4740
	cgattgccgg	gagtcccttt	cttctcatgt	caacgtgggt	acaagggagt	ctggcggggc	4800
	gacggcatca	tgcaaacacc	ctgccccatg	ggagcacaga	tcacccggaca	tgtagaaaaa	4860
20	ggtttccatga	ggatcgtggg	gcctaggacc	tgtagtaaca	cgtggcatgg	aacattcccc	4920
	attaacgcgt	acaccacggg	ccccctgcag	ccctccccgg	cgccaaatta	ttctaggggc	4980
	ctgtggcggg	tggtctgtga	ggagtacgtg	gaggttacgc	gggtggggga	tttccactac	5040
	gtgacgggca	tgaccaactga	caacgtaaag	tgcccggtgc	aggttccggc	ccccgaattc	5100
	ttcacagaag	tggtatgggt	gcgggtgcac	aggtacgctc	cagcgtgcaa	acccctccta	5160
	cgggagggag	tcacattcct	ggtcgggctc	aatcaatacc	tggttgggtc	acagctccca	5220
	tgcgggcccg	aaccggacgt	acttccatgc	tcacccgacc	ctccacatt	ctccacatt	5280
25	acggcggaga	cggctaagcg	taggtctggc	aggggatctc	ccccctcctt	ggccagctca	5340
	tcagctagcc	agctgtctgc	gccttccctt	aaggcaacat	gcactaccgc	tcagtactcc	5400
	ccggacgctg	acctcatcga	ggccaaacct	ctgtggcggc	aggagatggg	cgggaacatc	5460
	acccgcgtgg	agtcagaaaa	taaggtagta	atthtggact	cttccgagcc	gctccaagcg	5520
	gaggaggatg	agaggggaag	atccgttccg	gcggagatcc	tgccggaggtc	caggaaattc	5580
	cctcgagcga	tgcccatatg	ggcacgccc	gattacaacc	ctccactgtt	agagtcctgg	5640
30	aaggaccccg	actacgtccc	tccagtggta	cacgggtgtc	cattgccgcc	tgccaaggcc	5700
	cctccgatac	cacctccacg	gagggagagg	acggttgtcc	tgtcagaate	taccgtgtct	5760
	tctgccttgg	cggagctcgc	cacaaagacc	ttcggcagct	ccgaatcgct	ggcctccgac	5820
	agcggcacgg	caacggcctc	tcctgaccag	ccctccgacg	acggcgacgc	gggatccgac	5880
	gttgagtcgt	actcctccat	gccccccctt	gagggggagc	cgggggatcc	cgatctcagc	5940
	gacgggtctt	ggtctaccgt	aagcagggag	gctaagtggg	acgtcgtctg	ctgctcgatg	6000
	tcctacacat	ggacagggcg	cctgatcacg	ccatgcgctg	cggaggaaac	caagctgccc	6060
35	gtcaatgcac	tgagcaactc	tttgcctcgt	caccacaact	tggtctatgc	tacaacatct	6120
	cgcagcgcaa	gcctgcggca	gaagaaggct	acctttgaca	gactgcaggt	cctggacgac	6180
	cactaccggg	acgtgtctaa	ggagatgaag	gcgaaggcgt	ccacagttaa	ggctaaactt	6240
	ctatccgtgg	aggaagcctg	taagctgacg	ccccacatt	cggccagatc	taaatattgg	6300
	tatggggcaa	aggacgtccg	gaacctatcc	agcaaggccg	ttaaccacat	ccgctccgtg	6360
	tggaaggact	tgctgggaaga	cactgagaca	ccaattgaca	ccaccatcat	ggcaaaaaat	6420
	gaggttttct	gcgtccaacc	agaaagggg	ggccgcgaagc	cagctcgcct	tatcgtatct	6480
40	ccagatttgg	gggttcgtgt	gtgcgagaaa	atggcccttt	acgatgtggg	ctccaccttc	6540
	cctcaggccg	tgatgggctc	ttcatacgga	ttccaatact	ctcctggaca	gcgggtcgag	6600
	ttcctgggtg	atgcctggaa	agcgaagaaa	tgccctatgg	gcttcgcata	tgaccaaccg	6660
	tggtttgact	caacgggtcac	tgagaatgac	atccgtgttg	aggagtcaat	ctaccaatgt	6720
	tgtgacttgg	cccccgaa	cagacaggcc	ataaggctgc	tcacagagcg	gctttacatc	6780
	ggggggcccc	tgactaattc	taagggcgag	aactgcggct	atcgccggtg	ccgcgcgagc	6840
45	ggtgtactga	cgaccagctg	cggtaataac	ctcacatgtt	atttgaaggc	cgtgcggccc	6900
	tgctcagctg	cgaagctcca	ggactgcacg	atgctcgtat	cgggagacga	ccttgtcgtt	6960
	atctgtgaaa	gcgcggggac	ccaagaggac	gagggcgacc	tacgggcctt	cacggaggct	7020
	atgactagat	actctgcccc	ccctggggac	ccgccccaac	cagaatacga	cttggagtgg	7080
	ataacatcat	gctcctccaa	tgtgtcagtc	gcgcacgatg	catctggcaa	aaggggtgtac	7140
	tatctcacc	gtgacccccc	cacccccctt	gcgcgggctg	cgtgggagac	agctgacac	7200
	actccagtca	attcctggct	aggcaacatc	atcatgtatg	cgccccctt	gtgggcaagg	7260
50	atgatcctga	tgactcattt	cttctccatc	cttctagctc	aggaacaact	tgaaaaagcc	7320
	ctgattgtgc	agatctacgg	ggcctgttac	tccattgagc	cacttgacct	acctcagatc	7380
	attcaacgac	tccatggcct	tagcgcattt	tcactccata	gttactctcc	aggtgagatc	7440
	aataggggtg	cttcatgcct	caggaaactt	ggggtaccgc	ccttgcgagt	ctggagacat	7500

EP 1 043 399 A2

	cgggccagaa	gtgtccgcgc	taggctactg	tcccaggggg	ggagggctgc	cacttgtggc	7560
	aagtacctct	tcaactgggc	agtaaggacc	aagctcaaac	tcactccaat	cccggctgcg	7620
5	tcccagttgg	atztatccag	ctggttcggt	gctggttaca	gcgggggaga	catatatcac	7680
	agectgtctc	gtgcccgaac	ccgctgggtc	atgtgggtgc	tactcctact	ttctgtaggg	7740
	gtaggcacat	atctactccc	caaccgatga	acggggagct	aaacactcca	ggccaatagg	7800
	ccatcctggt	tttttccctt	tttttttttc	tttttttttt	tttttttttt	tttttttttt	7860
	ttttctcctt	tttttttcc	ttttctttcc	tttgggtggc	ccatcttagc		7920
	cctagtcacg	gctagctgtg	aaaggctcgt	gagccgcttg	actgcagaga	gtgctgatac	7980
10	tggcctctct	gcagatcaag	t				8001

15

20

25

30

35

40

45

50

55

<210> 7
 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus

5

<400> 7

10

15

20

25

30

35

40

45

50

55

```

gccagccccc gattgggggc gacactccac catagatcac tccccgtgta ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggctcg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccgc ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
ctcaaaagaa aaccaaacgt aacaccaacg ggccgcgcat gattgaacaa gatggattgc 420
acgcaggttc tcgggcccgt tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cagcaggggc gttccttgcg cagetgtgct ogacgttgte actgaagcgg 660
gaagggactg gctgctattg gggaagtgcc cggggcagga tctcctgtca tctcaccttg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctgcgcgacg 900
ccgaactggt cgccaggctc aaggcgcgca tgcccgacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgcttgcgc aatcatatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcgcc gaattgggctg accgcttccct cgtgtcttac ggtatcgccc 1140
ctcccgattc gcagcgcatc gccttctatc gccttcttga cgagttcttc tgagtttaa 1200
cagaccacaa cggtttccct ctacgggatc caattccgcc cctctccctc cccccccctt 1260
aacgttactg gccgaagccg cttggaataa ggccgggtgt cgtttgtcta tatgttatct 1320
tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctagggttct tccccctctc gccaaaggaa tgcaagggtc gttgaatgtc 1440
gtgaagggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccttt 1500
tgcaaggcagc ggaaccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaaggc ggcaaacacc cagtgcacag ttgtgagttg gatagttgtg 1620
gaaagagtca aatggtctct ctcaagcgta ttcaacaagg ggctgaagga tgcccagagg 1680
gtaccctcatt gtatgggata ggtctgggg cctcgggtga catgctttac atgtgtttag 1740
tcgaggttaa aaaaagctcta gggcccccca accacgggga cgtgggtttc ctttgaaaaa 1800
cacgataata ccatgggcac gaatcctaaa cctcaagaa aaaccaaacy taaccacaac 1860
cgccgcccac aggaagtcac gttcccgggc ggtggtcaga tctcgggtg agtttacctg 1920
ttgcgcgcga ggggccccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggtcg 1980
caacctcgtg gaaggcgaca acctatcccc aaggctcgcc agcccagagg tagggcctgg 2040
gctcagctcg ggtaccctcg gccctctat ggcaatgagg gcttgggggt ggcaggatgg 2100
ctcctgtcac cccgtggtcg tcggcctagt tggggcccca cggacccccg gcgtaggtcg 2160
cgcaatttgg gtaaggatcat cgataccctc acgtgcggct tcgcgatct catggggtac 2220
attcgcctcg tcggcgcccc cctagggggc gctgccaggg cectggcgca tggcgctcgg 2280
gttctggagg acggcggtgaa ctatgcaaca gggaaatctgc ccggttgctc cttttctatc 2340
ttccttttgg ctttctgtgc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400
gtatccggag tgtaccatgt cacgaacgac tgctccaaag caagcattgt gtaggggca 2460
gcggacatga tcatgcatac ccccggttgc gtgcccgtcg ttcgggagaa caactcctcc 2520
cgctgctggg tagcgtcac tcccacgctc gcggccagga acgctagcgt ccccaactacg 2580
acgatacgac gccatgtcga tttgctcgtt gggcgggctg ctctctgctc cgctatgtac 2640
gtgggagatc tctgcggatc tgttttcttc gtcgcccagc tgttcacctt ctgcctcgc 2700
cggcacgaga cagtacagga ctgcaattgc tcaatatatc ccggccacgt gacaggteac 2760
cgatatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgca 2820
ttactcgga tcccacaagc tgtcgtggat atggtggcgg gggcccatcg gggagtctta 2880
gcgggccttg cctactattc catggtgggg aactgggcta aggttctgat tgtgatgcta 2940
ctcttgcgg cgtttgacgg gggaaacctat gtgacagggg gacgatggc caaaaacacc 3000
ctcgggatta cgtccctctt ttcacccggg tcatcccaga aaatccagct tghtaaacacc 3060
aacggcagct ggacacatcaa caggactgcc ctgaactgca atgactccct caacactggg 3120
ttccttgctg cgctgttcta cgtgcacaa gttcaactcat ctggatgcc agagcgcatg 3180
gccagctgca gcccacatga cgcgttcgct cagggtggg ggcccatcac ttacaatgag 3240
tcacacagct cggacagag gcttattgt tggcactacg caccgcggcc gtgcggtatc 3300
gtaccgcgg cgaggtgtg tggtecagt tactgcttca ccccaagccc tgtcgtggtg 3360
gggacgacgg accggttcgg cgtccctacg tacagttggg gggagaatga gacggacgtg 3420
ctgcttctta acaacacggc gccgcgcaa ggcaactggt ttggtgttac atggatgaat 3480
agcaactgggt tcaccaagac gtgcgggggc ccccggtga acatcggggg gatcggcaat 3540

```

EP 1 043 399 A2

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

aaaaccttga cctgccccac ggactgcttc cggaagcacc ccgaggccac ttacaccaag 3600
 tgtggttcgg ggcttgggtt gacacccaga tgcttggtcc actaccata caggcttttg 3660
 cactaccctt gcaactgtcaa ctttaccatc ttcaagggtta ggatgtacgt ggggggagtg 3720
 gagcacaggc tcgaagccgc atgcaattgg actcaggagg agcgtttgta cctggaggac 3780
 agggacagat cagagcttag cccgctgctg ctgtctacaa cggagtgga ggtattgccc 3840
 tgttccctca ccacccatcc ggctctgtcc actggtttga tccatctcca tcagaacgtc 3900
 gtggacgtac aataccctga cggatatagg tcggcgggtt tctcctttgc aatcaaatgg 3960
 gagtattgcc tgttgcctct ccttctctct gcggaacgcgc gcgtctgtgc ctgcttgttg 4020
 atgatgtgct tgatagctca agctgagggc gccctagaga acctgggtgg cctcaacgcg 4080
 gcatccgtgg ccggggcgca tggcatctct tcttctctcg tgttctctcg tgcctgctgg 4140
 tacatcaaag gcaggctggg ccctggggcg gcataatgccc tctacggcgt atggccgcta 4200
 ctccctgtccc tgcctggcgtt accacacaga gcatacggca tggaccggga gatggcagca 4260
 tcgttgcagg gcgcgggttt cgttaggtctg atactcttga ccttgtcacc gcactataag 4320
 ctgttccctg ctaggctcat atgggtgggt caatatcttt cgggggggcc gcgatgccgt catcctctc 4380
 ttgcaagtgt ggcattcccc tccaccaga gctaattctt accatcacca aaatcttgc cgcataactc 4440
 acgtgcgcga tccaccaga ggtggtctca ggtggtata accaaagtgc cgtacttctg gcgcgcacac 4500
 ggggtcatct gggtcatgcat gctgttgccg gctgttgccg aaggttgctg ggggtcatta tgtccaaatg 4560
 gctctcatga agttggccgc actgacaggt acgtacgttt atgacatct caccctactg 4620
 ccggactggg ccacgcgggg cctacgagac ctgtcgggtg cagttgagcc cgtcgtcttc 4680
 tctgatatgg agaccaaggt tatcacttgg ggggcagaca ccgcggcggt tgggggacac 4740
 atcttggggc tggccgtctc cgcgcgagg gggagggaga tacatctggg accggcagac 4800
 agccttgaag ggcagggggt gcgcctatta cggcctactc ccaacagacg 4860
 cgaggcctac ttggctgcat ccatcactag ctacacaggc gggacaggaa ccaggtcgag 4920
 ggggaggtcc aagtgtcttc caccgcaaca caatctttcc tggcgacctg cgtcaatggc 4980
 gtgtgttgga ctgtctatca tgggtgcggc tcaaagacc ttgcggccc aaaggccca 5040
 atcacccaaa tgtacaccaa tgtggaccag gacctcgtcg gctggcaagc gcccccggg 5100
 gcgcgttctc tgacaccatg cactcgcggc agctcggacc tttacttggg cagcaggcat 5160
 gccgatgtca ttccggtgcg ccggcggggc gacagcaggg ggagcctact ctccccagg 5220
 cccgtctcct acttgaaggg ctcttcgggc ggtccaactg tctgcccctc ggggcatgct 5280
 gtgggcatct ttccggctgc cgtgtgcacc cgaggggttg cgaaggcggt ggaactttgt 5340
 cccgtcgagt ctatgaaac cactatgcgg tccccgggt tcacggacaa ctgcctccct 5400
 ccggccgtac cgcagacatt ccagggtggc catctacacg cccctactgg tagcggcaag 5460
 agcactaagg tgccggctgc gtatgcagcc caagggtata aggtgcttgt cctgaaccgg 5520
 tccgtgcccg caacctagg ttctggggcg tatatgtcta aggcacatgg tatcgacct 5580
 aacatcagaa ccggggtagg gacctacacc acgggtgccc ccatcacgta ctccacctat 5640
 ggcaagtctt ttgccgacgg tgggtgtctc gggggcgccct atgacatcat aatatgtgat 5700
 gagtgcact caactgactc gacctatc ctgggcatcg gcacagtcct ggaaccaagg 5760
 gagacggctg gaggcgact cgtcgtgctc gccaccgcta cgcctccggg atcgggtacc 5820
 gtgccacat caaacatcga ggaggtggct ctgtccagca ctggagaaa ccccttttat 5880
 ggcaaaagcca tccccatgc gacctcaag ggggggaggc acctcatttt ctgccattcc 5940
 aagaagaaat gtgatgagct cgcccggaag ctatccggcc tcggactcaa tgcgttagca 6000
 tattaccggg gccttgatgt atccgtcata ccaactagcg gagacgtcat tgcgttagca 6060
 acggacgctc taatgacggg ctttaccggc gatttcgact cagtgtatga ctgcaatata 6120
 tgtgtcacc agacagtcga cttacgcctg gaccgcacct tcaccattga gacgacgacc 6180
 gtgccacaag acgcggtgtc accctcgag cggcgaggca ggactggtag gggcaggatg 6240
 ggcatttaca ggtttgtgac tccaggagaa cggccctcgg gcatgttcga ttcctcggtt 6300
 ctgtgcgagt gctatgacgc gggctgtgct tggtagagc tcacgcccgc cgagacctca 6360
 gttaggttgc gggcttacct aaacacacca gggctgccc tctgccagga ccatctggag 6420
 ttctgggaga gcgtctttac aggcctcacc cacatagacg cccatttctt gtcccagact 6480
 aggcaggcag gagacaactt cccctacctg gtagcatacc aggtctacgg gtgcgccagg 6540
 gctcaggctc cactccatc atgtgggag gtctcatacg gctaaagcct gtaaggtact 6600
 acgctgcacg ggccaacgccc cctgctgtat aggtctggag ccgttcaaaa cgaggttact 6660
 accacacacc ccataaccaa atacatcatg gcatgcatgt cggctgacct ggaggtcgta 6720
 acgagcacct ggggtgctgg aggcggagtc ctagcagctc tggccgcgta ttgcctgaca 6780
 acaggcagcg tggctattgt gggcaggatc atcttgtccg gaaagccggc catcattccc 6840
 gacagggaag tcctttaccg ggagtctgat gagatggaag agtgccctc acacctccct 6900
 tacatcgaac agggaaatgca gctcgcgaa caattcaaac agaaggcaat cgggttgctg 6960
 caaacagcca ccaagcaagc ggaggtgct gctcccgtg tggaaatcaa gtggcgagacc 7020
 ctcaagcct tctggcgaa ctctgcctgg caaccccgcg atagcatcac tgatggcatt cacagcctct 7080
 atcaccagcc cgtcaccac ccaacatacc ctctgttta acatcctggg gggatgggtg 7140
 gccgccaac ttgctcctcc cagcgtgct tctgctttcg taggcggcg catcgctgga 7200
 gcgctgtttg gcagcatagg ccttgggaag gtgcttggg atattttggc aggttatgga 7260
 gcagggggtg caggcgcgct cgtggcctt aaggtcatga gcggcgagat gccctccacc 7320
 gaggacctg ttaacctact cctgctatc ctctcccctg gcgcctagt cgtcggggtc 7380
 7440
 7500

EP 1 043 399 A2

	gtgtgcgcag	cgatactgcg	tcggcacgtg	ggcccagggg	agggggctgt	gcagtggatg	7560
	aaccggctga	tagcgttcgc	ttcgcggggt	aaccacgtct	ccccacgca	ctatgtgtct	7620
	gagagcgacg	ctgcagcacg	tgtcactcag	atcctctcta	gtcttaccat	cactcagctg	7680
5	ctgaagagagc	ttcaccagtg	gatcaacgag	gactgtctcca	cgccatgctc	cggctcgtgg	7740
	ctaagagatg	tttgggattg	gatatgcacg	gtgttgactg	atttcaagac	ctggctccag	7800
	tccaagctcc	tgccgcgatt	gccgggagtc	cccttcttct	catgtcaacg	tgggtacaag	7860
	ggagtctggc	ggggcgacgg	catcatgcaa	accacctgcc	catgtggagc	acagatcacc	7920
	ggacatgtga	aaaacggttc	catgaggatc	gtggggccta	ggacctgtag	taacacgtgg	7980
	catggaacat	tccccattaa	cgcgtacacc	acgggccccct	gcacgccccct	cccgccgcca	8040
10	aattattcta	gggcgctgtg	gcgggtggct	gctgaggagt	acgtggagggt	tacgcgggtg	8100
	ggggatttcc	actacgtgac	gggcatgacc	actgacaacg	taaagtgcc	gtgtcaggtt	8160
	ccggcccccg	aattcttcac	agaagtggat	ggggtgcggt	tgcacaggta	cgctccagcg	8220
	tgcaaaccct	tcctacggga	ggaggtcaca	ttcctggctg	ggctcaatca	atacctgggt	8280
	gggtcacagc	tcccatgccc	gcccgaaacc	gacgtagcag	tgctcacttc	catgtcacc	8340
	gacccctccc	acattacggc	ggagacggct	aagcgtaggc	tggccagggg	atctcccccc	8400
15	tccttggcca	gctcatcagc	tagccagctg	tctgcgcctt	ccttgaaggc	aacatgcact	8460
	acccgtcatg	actccccgga	cgtgacctc	atcgaggcca	acctcctgtg	gcggcaggag	8520
	atgggcgggg	acatcacccg	cgtggagtca	gaaaataaag	tagtaatttt	ggactctttc	8580
	gagccgctcc	aagcggagga	ggatgagagg	gaagtatccg	ttccggcgga	gatcctgcgg	8640
	aggtccagga	aattccctcg	agcgtatgcc	atatgggcac	gcccggatta	caaccctcca	8700
	ctgttagagt	cctggaagga	cccggactac	gtccctccag	tggtaacagg	gtgtccattg	8760
20	ccgcctgcca	aggccctccc	gataccacct	ccacggaggg	agaggacggg	tgctcgtgca	8820
	gaatctaccg	tgtcttctgc	cttggcggag	ctcgccacaa	agaccttcgg	cagctccgaa	8880
	tcgtcggccg	tcgacagcgg	cacggcaacg	gcctctcctg	accagccctc	cgacgacggc	8940
	gacgcgggat	ccgacgttga	gtcgtactcc	tcctatgccc	cccttgaggg	ggagccgggg	9000
	gatcccgatc	tcagcgacgg	gtcttggctc	accgtaagcg	aggaggctag	tgaggacgtc	9060
	gtctgtgctc	cgatgtccta	catatggaca	ggcgccctga	tcacgccatg	cgctgcggag	9120
25	gaaaccaagc	tgcccgtcaa	tgaactgagc	aactctttgc	tccgtcacca	caacttggtc	9180
	tatgctacaa	catctcgtag	cgcgaagcct	cgagagaaga	aggtcacctt	tgacagactg	9240
	caggctcctg	acgaccacta	ccgggacgtg	ctcaaggaga	tgaaggcgaa	ggcgtccaca	9300
	gttaaggcta	aacttctatc	cgtggaggaa	gcctgttaag	tgacgcccc	acatttcggc	9360
	agatctaata	ttggctatgg	ggcaaaggac	gtccggaaac	tatccagcaa	ggccgttaac	9420
	cacateccgt	ccgtgtggaa	ggacttgctg	gaagacactg	agacaccaat	tgacaccacc	9480
30	atcatggcaa	aaaatgaggt	tttctgcgtc	caaccagaga	agggggggcc	caagccagct	9540
	cgccctatcg	tattcccaga	tttgggggtt	cgtgtgtgct	agaaaaatgg	cctttaagat	9600
	gtggtctcca	ccctccctca	ggcgtgatg	ggctcttcat	acggattcca	atactctcct	9660
	ggacagcggg	tcgagttcct	ggtgaatgcc	tggaaaagcg	agaaaatgcc	tatgggcttc	9720
	gcatatgaca	cccgtgtgtt	tgactcaacg	gtcactgaga	atgacatccg	tgttgaggag	9780
35	tcaatctacc	aatgttgtga	cttggccccc	gaagccagac	aggccataag	gtcgtctaca	9840
	gagcggcttt	acatcggggg	ccccctgact	aattctaaag	ggcagaactg	cggctatcgc	9900
	cgggtgcgcg	cgagcgggtg	actgacgacc	agctgcggta	ataccctcac	atgttatattg	9960
	aaggccgctg	cggcctgtcg	agctgcgaag	ctccaggact	gcacgatgct	cgatgcgga	10020
	gacgaccttg	tcgttatctg	tgaaaagcgc	gggacccaag	aggacgaggc	gagcctacgg	10080
	gccttcacgg	aggctatgac	tagatactct	gccccccctg	gggaccgcgc	caaaccagaa	10140
40	tacgacttgg	agttgataac	atcatgtctc	tccaatgtgt	cagtcgcgca	cgatgcactc	10200
	ggcaaaaagg	tgtactatct	caccctgac	cccaccaccc	cccttgccgc	ggctgcgtgg	10260
	gagacagcta	gacacactcc	agtcaattcc	tggctaggca	acatcatcat	gtatgcgccc	10320
	accttgtggg	caaggatgat	cctgatgact	catttcttct	ccatccttct	agctcaggaa	10380
	caacttgaaa	aagccctaga	ttgtcagatc	tacggggcct	gttactccat	tgagccactt	10440
	gacctacctc	agatcattca	acgactccat	ggccttagcg	cattttcact	ccatagtta	10500
45	tctccagggt	agatcaatag	ggtggcttca	tgcttcaggga	aacttggggg	accgcccttg	10560
	cgagtctgga	gacatcgggc	cagaagtgtc	cgcgctaggc	tactgtccca	ggggggggag	10620
	gctgccaact	gtggcaagta	cctcttcaac	tgggcagtaa	ggaccaagct	caaactcact	10680
	ccaatcccgg	ctcgttccca	gttggattta	tcagctgggt	tcgttgctgg	ttacagcggg	10740
	ggagacatat	atcacagcct	gtctcgtgcc	cgaacccgct	ggttcatgtg	gtgcctactc	10800
	ctactttctg	taggggtagg	catctatcta	ctccccaaac	gatgaacggg	gagctaaaca	10860
50	ctccaggcca	ataggccatc	ctgttttttt	cccttttttt	ttttcttttt	tttttttttt	10920
	tttttttttt	tttttttttc	tccttttttt	ttctcttttt	tttccctttt	tttccctttg	10980
	tggtctccat	ttagccctag	tcacggctag	ctgtgaaagg	tccgtgagcc	gcttgactgc	11040
	agagagtgtc	gatactggcc	tctctgcaga	tcaagt			11076

55

<210> 8
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

<400> 8
 gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
 ccccccctccc gggagagcca tagtggctcg cggaaccggt gagtacaccg gaattgccag 180
 gacgaccggg tcctttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtagt gcctgatagg 300
 gtgcttgccg gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
 ctcaagaaaa aaccaaacgt aacaccaacg ggcgcgcat gattgaacaa gatggattgc 420
 acgaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
 caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt 540
 ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
 cgtggctggc cagcagggcg gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
 gaaggggact gctgtatttg ggcgaaagtc cggggcagga tctcctgtca tctcaccttg 720
 ctctgcccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
 cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
 tggagcgctg tcttgcgat caggatgatc tggacgaaga gcacagggg ctccgcggag 900
 ccgaactgtt cgccaggctc aaggcgcgca tgcccagcgg cgaggatctc gtcgtgaccc 960
 atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
 actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtagata 1080
 ttgtgaaga gcttggcggc gaattggctg accgcttctc cgtgctttac ggtatcgccg 1140
 ctcccgatcc gcagcgcatc gccttctatc gccttcttga cgagttcttc tgagttttaa 1200
 cagaccacaa cggtttccct ctacgggatc caattccgcc cctctccctc cccccccct 1260
 aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
 tccaccatat tgcgctcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
 acgagcattc ctagggggtc ttcccctctc gccaaaggaa tgcaaggtct gttgaatgtc 1440
 gtgaaggaag cagtctctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
 tgcaggcagc ggaaccccc acctggcgac aggtgectct gcggccaaaa gccacgtgta 1560
 taagatacac ctgcaaaagg ggcacaaccc cagtgcacg ttgtgagttg gatagttgtg 1620
 gaaagagtca aatggctctc ctcaagcgtc ttcaacaagg ggctgaagga tgcccagaag 1680
 gtacccattt gtatgggata tgatctgggg cctcggtgca catgctttac atgtgtttag 1740
 tccagggttaa aaaacgtcta gggccccgga accacgggga cgtgggtttc ctttgaaaaa 1800
 cagcataata ccatggcgcc tatcagggc tactcccaac agacgcgagg cctacttggc 1860
 tgcacatcac ctacgttcac agggccggac aggaaccagg tcgaggggga ggtccaagtg 1920
 gtctccaccg caacacaatc tttctggcg acctgcgtca atggcggtg ttggactgtc 1980
 tatcatgtgt ccggtcaca gaccttggc ggcccaagg gcccaatcac ccaaatgtac 2040
 accaatgtgg accaggacct cgtcggtg caagcgcccc ccggggcgcg ttctttgaca 2100
 ccatgcacct gcggcagtc gcacctttac ttggtcacga ggcatgccga tgtcatctcg 2160
 gtgcgcccgc gggcgacag cagggggagc ctactctccc ccaggcccg etctacttg 2220
 aagggtctct cggcggtcc actgctctgc cctcgggg acgctgtggg catctttcgg 2280
 getgcggtgt gcacccgagg ggttgcaag gcggtggact ttgtaccgt cgagtctatg 2340
 gaaaccacta tgcggtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgcag 2400
 acattccagg tggcccatct acacgcccct actggtagcg gcaagagcac taagggtgcc 2460
 gctgcgtatg cagcccaagg gtataagggt cttgtcctga accggtccgt cgccgccacc 2520
 ctagggtttc gggcgatat gtctaaggca catgggtatc accctaacat cagaatcggg 2580
 gtaaggacca tcaccacggg tgcctccatc acgtactcca cctatggcaa gtttcttgcc 2640
 gacgggtggt gctctggggg cgcctatgac atcataatat gtgatgagtg ccactcaact 2700
 gactcgacca ctatcctggg catcggcaca gtcttgacc aagcggagac ggctggagcg 2760
 cgactcgtcg tgcctgccac cgtacgct cggggatcgg tcaccgtgcc acatccaaac 2820
 atcgaggagg tggctctgtc cagcactgga gaaatccct ttatggcaa agccatcccc 2880
 atcgagacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940
 gagctcgccg cgaagctgtc cggcctcgga ctcaatgtg tagcatatta ccggggcctt 3000
 gatgtatccg tcataccaac tagcggagac gtcatgtcg tagcaacgga cgctctaag 3060
 acgggcttta ccggtgactt cgactcagtg atcgactgca atacatgtgt caccagaca 3120
 gtcgacttca gctggaccc gaccttacc attgagacga cgaccgtgcc acaagacgcg 3180
 gtgtcacgct cgacgcggc aggcaggact ggtaggggca ggatgggcat ttacaggttt 3240
 gtgactccag gagaacggcc ctcgggcatg ttcgattcct cggttctgtg cgagtgtat 3300
 gacgggggtc gtgcttggtg cgaagtcacg ccgcccga cctcagttag gttgcgggct 3360
 tacctaaca caccaggtt gccgctctg caggaccatc tggagttctg ggagagcgtc 3420
 ttacagggc tcacccacat agacgccc atcttctccc agactaagca ggaggagac 3480
 aacttccct acctggtagc ataccaggct acggtgtgcg ccagggtca ggtccacct 3540

	ccatcgtggg	accaaagtgt	gaagtgtctc	atacggctaa	agcctacgct	gcacggggcca	3600
	acgcccctgc	tgtataggct	gggagccgtt	caaaacgagg	ttactaccac	acaccccata	3660
	accaaataca	tcatggcatg	catgtcggct	gacctggagg	tcgtcacgag	cacctgggtg	3720
5	ctggtaggcg	gagtcctagc	agctctggcc	gcgtattgcc	tgacaacagg	cagcgtggtc	3780
	attgtgggca	ggatcatctt	gtccggaaaag	ccggccatca	ttcccagacg	ggaaagtcctt	3840
	taccgggagt	tcgatgagat	ggaagagtgc	gcctcacacc	ttcccttacct	cgaacaggga	3900
	atgcagctcg	ccgaacaatt	caaacagaaag	gcaatcgggt	tgctgcatac	agccaccaag	3960
	caagcggagg	ctgctgctcc	cggtgtgaa	ttcaagtggc	ggaccatcga	agccttcttg	4020
	gcgaagcata	tgtggaattt	catcagcggg	atacaatatt	tagcaggctt	gtccactctg	4080
	cctggcaacc	ccgcgatagc	atcactgatg	gcattcacag	cctctatcac	cagccccctc	4140
10	accacccaac	ataccctcct	gtttaacatc	ctggggggat	gggtggccgc	ccaacttctg	4200
	cctcccagcg	ctgcttctgc	tttcgtaggc	gccggcatcg	ctggagcggc	tgttggcagc	4260
	atagcgcttg	ggaaggtgct	tgtggatatt	ttggcaggtt	atggagcagg	gggtggcagg	4320
	gcgctctcgc	cctttaagggt	catgagcggc	gagatgccct	ccaccgagga	cctgggttaac	4380
	ctactccctg	ctatcctctc	ccctggcgcc	ctagtctctg	gggtcgtgtg	cgcagcgata	4440
	ctgcgtcggc	acgtggggcc	aggggagggg	gctgtgcagt	ggatgaaccg	gctgatagcg	4500
15	ttcgcttcgc	gggtaacca	cgctctcccc	acgcactatg	tgctgagag	cgacgctgca	4560
	gcacgtgtca	ctcagatcct	ctctagtctt	accatcactc	agctgctgaa	gaggcttcac	4620
	cagtgatca	acgaggactg	ctccacgcca	tgctccggct	cggtggctaag	agatgttttg	4680
	gattggatat	gcacgggtgt	gactgatttc	aagacctggc	tcacgtccaa	gctcctgccc	4740
	cgattgccc	gagtcctctt	cttctcatgt	caacgtgggt	acaagggagt	ctggcggggc	4800
	gacggcatca	tgcaaacacc	ctgcccattg	ggggcacaga	tcacccggaca	tgtgaaaaaac	4860
	ggttccatga	ggatcgtggg	gcctaggacc	tgtagtaaca	cggtggcatgg	aacattcccc	4920
20	attaaacgct	acaccacggg	cccttgacag	ccctccccgg	cgccaaatta	ttctaggggc	4980
	ctgtggcggg	tggtctgtga	ggagtacgtg	gaggttacgc	gggtggggga	tttccactac	5040
	gtgacgggca	tgaccactga	cgacgtaaag	tgcccgtgtc	aggttccggc	ccccgaattc	5100
	ttcacagaag	tggtgggggt	gcggttgac	aggtacgctc	cagcgtgcaa	acccctccta	5160
	cgggaggagg	tcacattcct	ggtcggggct	aatcaatacc	tggttgggtc	acagctccca	5220
	tgcgagccctg	aaccggatgt	agcagtgtct	acttccatgc	tcaccgaccc	ctcccacatt	5280
25	acggcggaga	cggttaagcg	taggtctggc	aggggatctc	ctcccccttt	ggccagctca	5340
	tcagctagcc	agctgtctgc	gecttctctg	aaggcaacat	gcactaccgg	tcagtactcc	5400
	ccggacgcctg	acctcatcga	cgccaacctc	ctgtggcgcc	aggagatggg	cgggaaacac	5460
	acccgcgtgg	agtcaaaaa	taaggtagta	attttggaat	ctttcgagcc	gctccaagcg	5520
	gaggaggatg	agaggggaag	atccgttccg	gcggagatcc	tgccgagggt	caggaaattc	5580
	cctcgagcga	tgcccatatg	ggcacgcccc	gattacaacc	ctccactggt	agagtctctg	5640
	aaggaccccg	actacgtccc	tccagtggta	cacgggtgtc	cattgcccgc	tgccaaagcc	5700
30	cctccgatac	caccttcacg	gaggaagagg	acggttgttc	tgctcagaatc	taccgtgtct	5760
	tctgcccctg	cggagctcgc	cacagagacc	ttcggcagct	ccgaatcgct	ggccgtctgac	5820
	agcggcacgg	caacggcctc	tcctgaccag	ccctccgacg	acggcgacgc	gggatccgac	5880
	ggttagtctg	actcctccat	gccccctctt	gagggggagc	cgggggatcc	cgactctcag	5940
	gacgggtctt	ggtctaccgt	aagcgaggag	gctagtggag	acgtcgtctg	ctgctcgatg	6000
	tcctacacat	ggacaggcgc	cctgatcagc	ccatgcgctg	cggaggaaac	caagctgccc	6060
35	atcaatgcac	tgagcaactc	tttctcctgt	caccacaact	tggtctatgc	tacaacatct	6120
	cgcagcgcaa	gcctgcccga	gaagaaggtc	acctttgaca	gactgcaggt	cctggacgac	6180
	cactaccggg	acgtgtctca	ggagatgaag	gcgaaggcgt	ccacagttaa	ggctaaactt	6240
	ctatccgtgg	aggaagcctg	taagctgacg	ccccacatt	cggccagatc	taaatttggc	6300
	tatggggcaa	aggacgtccg	gaacctatcc	agcaaggccg	ttaaccacat	ccgctccgtg	6360
	tggaaggact	tgctggaaga	cactgagaca	ccaattgaca	ccacatcat	ggcaaaaaat	6420
	gaggttttct	gcgtccaacc	agagaagggg	ggccgcaagc	cagctcgccct	tatcgtattc	6480
40	ccagatttgg	gggttcgtgt	gtgcgagaaa	atggcccttt	acgatgtggt	ctccaccctc	6540
	cctcaggccg	tgatgggctc	ttcatagcga	ttccaatact	ctcctggaca	gcgggtcgag	6600
	ttcctggtga	atgcctggaa	agcgaagaaa	tgccctatgg	gcttcgcata	tgacacccgc	6660
	tgttttgact	caacgggtcac	tgagaatgac	atccgtgttg	aggagtcaat	ctaccaatgt	6720
	tgtgactctg	cccccgaaag	cagacagggc	ataaggtcgc	tcacagagcg	gctttacatc	6780
	gggggcccc	tgactaattc	taaagggcag	aactgcggct	atcgccgggtg	ccgcgcgagc	6840
45	ggtgtactga	cgaccagctg	cggttaatacc	ctcacatggt	acttgaaggc	cgctgcggcc	6900
	tgctcagctg	cgaagctcca	ggaactgcag	atgctcgtat	gcggagacga	ccttgtcgtt	6960
	atctgtgaaa	gcgcggggac	ccaagaggac	gagggcagcc	taaggggcctt	cacggaggct	7020
	atgactagat	actctgcccc	ccctggggac	ccgcccacac	cagaatacga	cttggaagttg	7080
	ataacatcat	gctcctccaa	tgtntcagtc	gcgcacgatg	catctggcaa	aaggggtgtac	7140
	tatctcacc	gtgaccccac	caccccctt	gcgcgggctg	cgctgggagac	agctagacac	7200
	actccagtca	attcctggct	aggcaacatc	atcatgtatg	cgcccacctt	gtgggcaagg	7260
50	atgatcctga	tgactcattt	cttctccatc	cttctagctc	aggaacaact	tgaaaaagcc	7320
	ctagattgtc	agatctacgg	ggcctgttac	tccattgagc	cacttgacct	acctcagatc	7380
	attcaacgac	tccatggcct	tagcgcatct	tcactccata	gttactctcc	aggtgagatc	7440
	aatagggtgg	cttcatgctt	caggaaactt	ggggtaccgc	ccttgccgag	ctggagacat	7500

EP 1 043 399 A2

	cgggccagaa	gtgtccgcgc	taggctactg	ttccaggggg	ggagggctgc	cacttggtggc	7560
	aagtacctct	tcaactgggc	agtaaggacc	aagctcaaac	tcactccaat	cccggctgcy	7620
5	ttccagttgg	atztatccag	ctggttcggt	gctggttaca	gcgggggaga	catatatcac	7680
	agcctgtctc	gtgcccgacc	ccgctgggtc	atgtgggtgc	tactcctact	ttctgtaggg	7740
	gtaggcatct	atctactccc	caaccgatga	acggggagct	aaacactcca	ggccaatagg	7800
	ccatcctgtt	tttttccctt	tttttttttc	tttttttttt	tttttttttt	tttttttttt	7860
	ttttctcctt	tttttttcc	cttttttttc	ttttctttcc	tttgggtggc	ccatcttagc	7920
	cctagtcacg	gctagctgtg	aaaggctccg	gagccgcttg	actgcagaga	gtgctgatac	7980
10	tggcctctct	gcagatcaag	t				8001

15

20

25

30

35

40

45

50

55

EP 1 043 399 A2

<210> 9
<211> 11076
<212> DNA
<213> Hepatitis C Virus

5

```

<400> 9
gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctcccc gggagagcca tagtggtctg cggaaccggg gagtacaccg gaattgccag 180
gacgaccggg tccttttctg gatcaacccg ctcaatgcct ggagatttgg gcgtgcccc 240
gagagactgc tagccagta gtgttgggtc gcgaaaggcc ttgtgttact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaaagaaa aaccaaactg aacaccaacg ggccgcgccat gattgaacaa gatggattgc 420
acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cagcagggg gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaaggagactg gctgctattg ggcgaagtgc cggggcaggga tctcctgtca tctcaccttg 720
ctcctgcgga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgtcgat caggatgac tggacgaaga gcatcagggg ctgcgcccag 900
ccgaactggt cgccaggtcc aaggcgcgca tgcccagcgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgcttgcg aatcatggtg tggaaaatgg ccgcttttct ggattatog 1020
actgtggccg gctgggtgtg gcggacgcgt atcaggacat agcgttggct acccgtgata 1080
ttgtgaaga gcttggcggc gaatgggtg accgcttctt cgtgctttac ggtagcgc 1140
ctcccgatcc gcagcgcatc gccttctatc gccttcttga cgagtctctc tgagtttaa 1200
cagaccacaa cgggttccct cttagcggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg ctggaataaa ggccggtgtg cgtttgtcta tatgttattt 1320
tccaccatat tgcgtctttt tggcaatgtg agggcccgga aacctggccc tgtcttctg 1380
acgagcattc ctagggttct tccccctctc gccaaaggaa tgcaaggctc gttgaatgtc 1440
gtgaagggaag cagttctctc ggaagcttct tgaagacaaa caacgtctgt agcagccctt 1500
tgccagcgag ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc gccacaaccc cagtgcacg ttgtgagttg gatagttgtg 1620
gaaagtgcca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtaccccatc gtatgggatc tgatctgggg cctcggtgca catgctttac atgtgttag 1740
tcgaggttaa aaaacgtcta ggcccccgga accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccatgggcac gaatcctaaa cctcaayaa aaaccaaacy taacaccaac 1860
cgccgcccc caggacgtcaa ggtcccgggc ggtggtcaga tctcggtgg agtttacctg 1920
ttgccgcgca ggggccccag gttgggtgtg cgccgcgacta ggaagacttc cgagcggtcg 1980
caacctcggt gaaggcgaca acctatcccc aaggctcgcc agcccgaggg tagggcctgg 2040
gctcagcccc ggtaccctcg gccctcttat ggcaatgagg gcttggggtg ggcaggatgg 2100
ctcctgtcac cccgtggctc tgggccccca cggacccccg gcgtaggtcg 2160
cgcaatttgg gtaaggtcat cgataccctc acgtgcggct tcgcccgtat catggggtac 2220
attccgctcg tcggcgcccc gctggcgggc cctggcgcca tggcgctcgg 2280
gttctggagg acggcgtgaa ctatgcaaca gggaatctgc ccggttgctc ctttctctatc 2340
ttccttttgg ctttgcgtgc ctgtttgacc atcccagctt ccgcttatga agtcgcaac 2400
gtatccggag tgtaccatgt cacgaacgac tgctccaacg caagcattgt gtatgaggca 2460
gaggacatga tcatgcatac ccccggtgac gtgcccgtcg ttcgggagaa caactcctcc 2520
cgctgctggg tagcgtctac tcccacgctc gccgcccagg acgctagcgt cccactacg 2580
acgatacgac gccatgtcga tttgctcggt gggcgcgctg ctctctgctc cgctatgtac 2640
gtgggagatc tctgcggatc tgttttcttc gtcgcccagc tgttcaacct ctcgctcgc 2700
cggcacgaga cagtaaggga ctgcaattgc tcaatatatc ccggccaagt gacaggtcac 2760
cgtatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgag 2820
ttactccgga tcccacaagc tgtcgtggat atgggtggcg gggccattg gggagctcta 2880
gcgggccttg cctactattc catgttgggg aactgggcta aggttctgat tgtgatgcta 2940
ctctttgccc gcgttgacgg gggaaacctat gtgacagggg ggacgatggc caaaaaacac 3000
ctcgggatta cgtccctctt ttaccccggg tcatcccaga aaatccagct tgtaaaacac 3060
aacgcagctt ggcacatcaa caggactgcc ctgaactgca atgactccct caacactggg 3120
ttccttgctg cgctgttcta cgtgcacaag ttcaactcat ctggatgcc agagcgcatg 3180
gccagctgca gcccacatga cgcgttcgct cagggtgggg ggcccatcac ttacaatgag 3240
tcacacagct cggaccagag gccttattgt tggcactacg caccocggcc gtgcggtatc 3300
gtaccgcggc cgcaggtgtg tggccagtgt tactgcttca ccccaagccc tgcgtgtgtg 3360
gggacgacgg accggttcgg cgtccctacg tacagtgggg gggagaatga gacggacgtg 3420
ctgctcttta acaacacggc gcccgcgcaa ggcaactggt ttgctgtgac atggatgaat 3480
agcactgggt tcaccaagac gtgcgggggc ccccgtgta acatcggggg gatcggaat 3540

```

55

5

10

15

20

25

30

35

40

45

50

55

```

aaaaccttga cctgccccac ggactgcttc cggaagcacc ccgaggccac ttacaccaag 3600
tgtgtgttcgg ggccttgggt gacacccaga tgccttgggtc actaccata caggctttgg 3660
cactaccctt gcaactgtcaa ctttaccatc ttcaagggtta ggatgtacgt ggggggagtg 3720
gagcacaggc tcgaagccgc atgcaattgg actcgaggag agcgttgtaa cctggaggag 3780
agggacagat cagagcttag ccgcgtgctg ctgtctacaa cggagtggca ggtattgccc 3840
tgttctctta ccaccctacc ggctctgtcc actggtttga tccatctcca tcagaacgtc 3900
gtggacgtac aatacctgta cggatagagg tcggcggttg tctcctttgc aatcaaatgg 3960
gagtatgtcc tgttgctctt ccttcttctg gcggacgcgc gcgtctgtgc ctgcttgtgg 4020
atgatgtgct tgatagctca agctgaggcc gcctagaga acctgggtgg cctcaacgcg 4080
gcatccgtgg ccggggcgca tggcattctc tccttctctg tgttctctg tgcctgctgg 4140
tacatcaagg gcaggctggt ccctggggcg gcatatgccc tctacggcgt atggccgcta 4200
ctcctgctcc tgctggcggt accaccacga gcatacgcca tggaccggga gatggcagca 4260
tcgtgctggg gcgcgggttt cgtaggtctg atactcttga ccttgtcacc gcactataag 4320
ctgttctctg ctaggctcat atggtggtta caatatttta tcaccagggc cgaggcacac 4380
ttgcaagtgt ggatcccccc cctcaacgtt cggggggggc gcgatgccgt catcctctc 4440
acgtgcgcga tccaccaga gctaactctt accatcacca aaatcttct cgcctactc 4500
ggctccactc tgggtctcca ggtggtgata accaaagtgc cgtacttctg gcgcgcacac 4560
gggctcattc gtgcatgcat ggtgtgctgg aaggttgcgt ggggtcatta tgtccaaatg 4620
gctctcatga agttggccgc actgacaggt acgtacgttt atgaccatct cccccactg 4680
cgggaactgg cccacgcggg cctacgagac cttgcgggtg cagttgagcc cgtcgtctc 4740
tctgatattg agaccaaggt tatcacctgg ggggcagaca ccgcggcgtg tggggacatc 4800
atccttgggg ccgccgtctc cgcccgaggg gggagggaga tacatctggg accggcagac 4860
agccttccag ggcagggggt cgcactctc gcgcctatta cggcctactc ccaacagacg 4920
cgaggcctac ttggctgcat catcactagc ctcacaggcc gggacaggaa ccaggtcgag 4980
ggggaggttc aagtgtctc caccgcaaca caatctttcc tggcgacctg cgtcaatggc 5040
ccgttcgaag ctgtctatca tggtgccggc tcaaagaccc ttgccggccc aaagggccca 5100
atcacccaaa tgtacaccaa tgtggaccag gacctcgtcg gctggcaagc gccccccggg 5160
gcgcgttctc tgacaccatg cacctgcggc gacagcaggg ggagcctact ctccccagg 5280
gctgtcttgc acttgaaggg ctcttcgggc gtaggggttg cgaaggcggg ggactttgta 5400
ccgtcgagat ctatgaaac cactatgcgg tcccgggtct tcacggacaa ctgcctcccc 5460
ccggccgtac cgcagacatt ccaggtggcc catctacacg ccctactagg tagcggcaag 5520
agcataaagg tgccggctgc gtatgcagcc caaggtgata aggtgcttgt cctgaacccg 5580
tccgtcgccg ccaccctagg ttccggggcg tatatgtcta aggcacatgg tatcgacct 5640
aacatcagaa tcggggtaag gaccatcacc acgggtgccc ccatcacgta ctccacctat 5700
ggcaagtttc ttgccgacgg tgggtgctct gggggcgccct atgacatcat aatatgtgat 5760
gagtgccact caactgactc gaccactatc ctgggcatcg gcacagtcct ggaccaagcg 5820
gagacggctg gagcgcgact cgtcgtgctc gccaccgcta cgcctccggg atcggctcac 5880
gtgccacatc caaacatcga ggaggtggct ctgtccagca ctggagaaat ccccttttat 5940
ggcaaaagcca tcccacatga gaccatcaag ggggggaggg acctcathtt ctgccattcc 6000
aagaagaaat gtgatgagct cgcgcgcaag ctgtccggcc tcggactcaa tgcgtgtagca 6060
tattaccggg gccctgatgt cctcgtcata ccaactagcg gagacgtcat tgcgttagca 6120
acggacgctc taatgacggg ctttacgggt gacttcgact cagtgatcga ctgcaatata 6180
tgtgtcaccc agacagtcga cttcagcctg gacccgacct tcaccattga gacgacgacc 6240
gtgccacaag acgcggtgct acgctcgca cggcgaggca ggaactggtag gggcaggatg 6300
ggcatttaca ggtttgtgac tccaggagaa cggccctcgg gcatgttcga ttccctcggt 6360
ctgtgcgagt gctatgacgc gggctgtgct tggtaacgag tcaegccgc cgagacctca 6420
gttaggttgc gggcttacct aaacacacca gggttgcccg tctgccaagg ccatctggag 6480
ttctgggaga gcgtctttac aggcctcacc cacatagacg cccatttctt gtcccagact 6540
aagcaggcag gagacaactt cccctacctg gtagcatacc aggctacggg gtgcgcagg 6600
gctcaggctc cacctccatc gtgggaccaa atgtggaagt gtctcatacg gctaaaagcct 6660
acgctgcacg ggccaacgcc cctgctgtat aggcctggag ccgttcaaaa cgaggttact 6720
accacacacc ccataaccaa atacatcatg gcatgcatgt cggctgacct ggaggtcgct 6780
acgagcacct ggggtgctgg ctacggagtc ctacgagctc tggccgcgta ttgcctgaca 6840
acaggcagcg tggctattgt gggcaggatc atcttgcctg gaaagccggc catcattccc 6900
gacaggggag tcctttaccg ggaagtccgat gagatggaag agtgccctc acacctccct 6960
tacatcgaac agggaaatgc gctcggcgaa caattcaaac agaagcaat cgggttgctg 7020
caaacagcca ccaagcaagc ggaggctgct gctcccgtgg tggaaatcaa gtggcgagc 7080
atcgaagcct tctggcgcaa gcataatgtg aatttcatca gcgggataca atatttagca 7140
ggcttgcctc ctctgcctgg caaccccgcg atagcatcac tgatggcatt cacagcctct 7200
atcacagacc cgctcaccac ccaacatacc ctctgttta acatctctgg gggatgggtg 7260
gccgccaac ttgctcctcc cagcgtgctc tctgctttcg taggcgcgg catcgctgga 7320
gcgctgtgtg ccttgggaag gtgcttggg atattttggc aggttatgga 7380
gcagggggtg caggcgctc cgtggccttt aaggtcatga gcggcgagat gccctccacc 7440
gaggacctg ttaacctact ccttgcctac ctctcccctg gcgcccctag cgtcggggtc 7500

```

5 gtgtgcgag cgatactgcg tcggcacgtg ggcccagggg agggggctgt gcagtggatg 7560
 aaccggctga tagcgttcgc ttccgsgggg aaccacgtct cccccacgca ctatgtgctt 7620
 gagagcgacg ctgcagcacg tgtcactcag atcctctcta gtcttaccat cactcagctg 7680
 ctgaagaggc ttaccagtg gatcaacgag gactgctcca cgccatgctc cggctcgtgg 7740
 ctaagagatg tttgggattg gatatgcacg gtgttgactg atttcaagac ctggtccag 7800
 tccaagctcc tgccgcgatt gccgggagtc cccttcttct catgtcaacg tgggtacaag 7860
 ggagtctggc ggggcgacgg catcatgcaa accacctgcc catgtggggc acagatcacc 7920
 ggacatgtga aaaacggttc catgaggatc gtggggccta ggacctgtag taacacgtgg 7980
 catggaacat tccccattaa cgcgtacacc acggggccct gcacgccctc cccggcgcca 8040
 10 aattattcta gggcgctgtg gcgggtggct gctgaggagt acgtggagggt tacgcgggtg 8100
 ggggatttcc actacgtgac gggcatgacc actgacgacg taaagtgcc gtgtcagggt 8160
 ccggcccccg aattcttcac agaagtggat ggggtgcggt tgcacaggta cgctccagcg 8220
 tgcaaacccc tccatcgga ggaggtcaca ttcttggtcg ggctcaatca atacctggtt 8280
 ggggtcacagc tccatcgga gccgaaccg gatgtagcag tgctcacttc catgtcacc 8340
 15 gacccctccc acattacggc ggagacggct aagcgtaggc tggccagggg atctctccc 8400
 cccttgggca gctcatcagc tagcagctg tctgcgcctt ccttgaaggc aacatgcact 8460
 acccgtcatg actccccgga cgtgacctc atcgaggcca acctctgtg gccgcaggag 8520
 atgggcggga acatcacccg cgtggagtca gaaaataagg tagtaatttt ggactcttcc 8580
 gagccgctcc aagcggagga ggatgagagg gaagtatccg tccggcgga gatcctgcgg 8640
 aggtccagga aattccctcg agcgtgccc atatgggac gcccgatta caacctcca 8700
 ctgttagagt cctggaagga cccggactac tccctccag tggtaacagg gtgtccattg 8760
 20 ccgcctgcca aggcctccc gataccacct tcacggagga agaggacggt tgtcctgtca 8820
 gaatctaccg tgtctctgc cttggcgag ctgccacag agaccttcgg cagctccgaa 8880
 tcgtcggccg tcgacagcgg caccgcaacg gcctctcctg accagccctc cgacgacggc 8940
 gacgcgggat ccgacgttga tctgtactcc tccatgcccc cccttgaggg ggagccgggg 9000
 gatcccgatc tcagcgacgg gtcttggtct accgtaagcg aggaggctag tgaggacgtc 9060
 gtctgctgct cgatgtccta cacatggaca ggcgcctga tcacgccatg cgctgcggag 9120
 25 gaaaccaagc tgcccatcaa tgcactgagc aactctttgc tccgtacca caactgggtc 9180
 tatgtctaca catctcgag cgcgaacctg cggcagaaga aggtcacctt tgacagactg 9240
 caggctcctg acgaccacta ccgggacgtg ctcaaggaga tgaaggcgaa ggcgtccaca 9300
 gtttaaggcta aacttctatc cgtggaggaa gcctgttaagc tgacgcccc acattcggcc 9360
 agatctaaat ttggctatgg ggcaaaggac gtccggaacc tatccagcaa ggcgttaac 9420
 30 cacatccgct ccgtgtggaa ggacttctg gaagacactg agacaccaat tgacaccacc 9480
 atcatggcaa aaaaagaggt tttctgcgtc caaccagaga agggggggccg caagccagct 9540
 cgccctatcg tattccaga tttgggggtt cgtgtgtgag agaaaaatggc cctttacgat 9600
 gtgggtctcca cctccctca ggcgtgatg ggctcttcat acggattcca atactctct 9660
 ggacagcggg tcgagttcct ggtgaatgcc tggaaagcga agaaatgcc tatgggcttc 9720
 gcatatgaca cccgtgttt tgactcaacg gtcaactgaga atgacatccg tgttgaggag 9780
 35 tcaatctacc aatgttctga ttggcccc aggccagac aggccataag gtcgctcaca 9840
 gagcggcttt acatcggggg cccctgact aattctaaag ggcagaactg cggctatcgc 9900
 cgggtgccgcg cgagcgggtg actgacgacc agctgcggtg atacctcac atgttacttg 9960
 aaggccgctg cggcctgtcg agctgcgaag ctccaggact gcacgatgct cgtatgcgga 10020
 gacgaccttg tcgttatctg tgaaagcgcg gggacccaag aggacgaggc gagcctaagg 10080
 gccttcacgg aggtatgac tagatactct gccccccctg gggacccgcc caaacagaa 10140
 40 tacgacttgg agttgataac atcatgctcc tccaatgtgt cagtcgcgca cgatgcattc 10200
 ggcaaaaggg tgtactatct caccgtgac cccaccaccc cccttgcgcg ggctgcgtgg 10260
 gagacagcta gacacactcc agtcaattcc tggctaggca acatcatcat gtatgcgccc 10320
 accttgtggg caaggatgat cctgatgact catctctctc ccatcctctc agctcaggaa 10380
 caacttgaag aagccctaga ttgtcagatc tacggggcct gttactccat tgagccactt 10440
 gacctacctc agatcattca acgactccat ggccttagcg catttctact ccatagtac 10500
 45 tctccaggtg agatcaatag ggtggcttca tgcctcagga aacttggggg accgcccttg 10560
 cgagtctgga gacatcgggc cagaagtgtc cgcgctaggc tactgtccca gggggggagg 10620
 gctgccactt gtggcaagta cctcttcaac tgggcagtaa ggaccaagct caaactcact 10680
 ccaatcccgg ctgcgtccca gttggattta tccagctggg tcggtgctgg ttacagcggg 10740
 ggagacatat atcacgcct gtctcggtcc cgaccccgct ggttcatgtg gtgcctactc 10800
 ctactttctg taggggtagg catctatcta cctcccaacc gatgaacggg gagctaaaca 10860
 50 ctccaggcca ataggccatc ctgttttttt cccttttttt ttttcttttt ttttcttttt 10920
 ttttcttttt ttttcttttt tctctttttt ttctcttttt tttctttttt tttctttttg 10980
 tggtccatc ttagccctag tcacggctag ctgtgaaagg tccgtgagcc gcttgactgc 11040
 agagagtgtc gatactggcc tctctgcaga tcaagt 11076

<210> 10
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

5

```

<400> 10
gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgcgtgag cctccaggac 120
ccccctccc gggagagcca tagtggctcg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaacccg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaggcc ttgtggtagt gcctgataag 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
ctcaaaagaa aaccaaactg aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcagggtc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cagcaggggc gttcccttgc cagctgtgct cgacgttgtc actgaagcgg 660
gaaggggact gctgtattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgca gcggtgtcat acgcttgatc 780
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagcccg tcttgcgat caggatgac tggaacgaaga gcatcagggg ctgcgcgcag 900
ccgaactggt cgccaggctc aaggcgcgca tgcccgacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggcgg gctgggtgtg gcggaaccgt atcaggacat agcgttggct acccgtaga 1080
ttgtgaaga gcttgccggc gaatgggctg accgcttct cgtgctttac ggtatcgccg 1140
ctcccgatc gcagcgcatc gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200
cagaccacaa cggtttccct ctacgggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccgggtgt cgtttgtcta tatgttatt 1320
tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctagggggtc tccccctct gccaaaggaa tgcaaggtct gttgaatgtc 1440
gtgaaggaag cagtctctct ggaagcttct tgaagacaaa caacgtctgt agcgacctt 1500
tgcaaggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc ggcaaacacc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtc aatggctctc ctcaagcgt a tcaacaagg ggctgaagga tgcccagaag 1680
tcaccccatc gtatgggata tgatctgggg cctcggtgca catgctttac atgtgttag 1740
tcgaggttaa aaaaagtcta ggcccccg aaccagggga cgtggtttcc ctttgaaaaa 1800
cacgataata ccatggcgcc tattacggcc tactcccaac agacggaggg cctacttggc 1860
tgcatcata ctacgtcac aggcggggac aggaaccagg tcgaggggga ggtccaagt 1920
gtctccaccg caacacaatc tttcctggcg acctgcgtca atggcggtg ttggactgtc 1980
tatcatggtg ccggtcctaa gaccttgcc ggcccaagg gcccaatcac ccaaatgtac 2040
accaatgtgg accaggacct cgtcggtgg caagcgcccc ccggggcgcg ttccttgaca 2100
ccatgcacct gcggcagctc ggacctttac ttggtcacga ggcatgccga tgtcatccg 2160
gtgcgcccgc ggggcgacag cagggggagc ctactctccc ccaggcccg ctctacttg 2220
aagggtctct cgggcgggtc actgctctgc cctcggggc acgctgtggg catctttcgg 2280
gctgcgtgtg gcaccgaggg ggttgcaagg gcggtggact ttgtaccgt cagttctatg 2340
ggaaccacta tgcgggtccc ggtcttcacg gacaactcgt cccctccggc cgtaccgag 2400
acattccagg tggcccatct acacgcccct actggtagcg gcaagagcac taagggtggc 2460
gctgcgtatg caggccaagg gtataagggt cttgtcctga acccgtcgt cgccgccacc 2520
ctaggtttcg gggcgatatat gtctaaggca catggtatcg accctaaccat cagaatcggg 2580
gtaaggacca tcaccacggg tgcccccatc acgtactcca cctatggcaa gtttcttgcc 2640
gacggtgtgt gctctggggg cgccatgac atcataatat gtgagtagtg ccaactcaat 2700
gactcgacca ctatcctggg catcggcaca gtcttgacc aagcggagac ggctggagcg 2760
cgactcgtcg tgctcgccac cgctacgctc ccgggatcgg tcaccgtgcc acatccaaa 2820
atcgaggagg tggctctgtc cagcactgga gaaatccct tttatggcaa agccatcccc 2880
atcgagacca tcaagggggg gaggcacctc atttctgccc attccaaga gaaatgtgat 2940
gagctcgccg cgaagctgtc cggcctcgga ctcaatgctg tagcatatta ccggggcctt 3000
gatgtatccg tcataccac tagcggagac gtcatgtctg tagcaacgga cgtctaatg 3060
acgggtctta ccggcgactt cgactcagtg atcgactgca atacatgtgt caccagaca 3120
gtcgacttca gcctggaccg gaccttcacc attgagacga cgaccgtgcc acaagacggc 3180
gtgtcacgct cgacggcg aggcaggact ggtaggggca ggatgggcat ttacaggtt 3240
gtgactccag gagaacggcc ctccggcatg ttcgattcct cggttctgtg cgagtgetat 3300
gacgggggtc gtgcttggtc cgagctcag cccgcccaga cctcagttag gttgcccgt 3360
tacctaaca caccaggtt gccgctcgc caggaccatc tggagttctg ggagagcgtc 3420
tttacaggcc tcaccacat agacgccc atcttgtccc agactaagca ggcaggagac 3480
aacttcccct acctggtagc ataccaggct acggtgtcgc ccagggtca ggctccacct 3540

```

55

	ccatcgtagg	acccaaalg	gaagtgtctc	atacggctaa	agcctacgct	gcacggggcca	3600
	acgccccctg	tgatataggc	gggagccgtt	caaaaacgagg	ttactaccac	acaccccata	3660
	accaaataca	tcatggcatg	catgtcggct	gaacctggagg	tcgtcacgag	cacctgggtg	3720
5	ctggtagggc	gagtcctagc	agctctggcc	gcgtattgcc	tgacaacagg	cagcgtgggtc	3780
	attgtgggca	ggatcatctt	gtccggaaa	ccggccatca	ttcccgacag	ggaagtcctt	3840
	taccgggagt	tcgatgagat	ggaagagtgc	gcctcacacc	tcctttacat	cgaacaggga	3900
	atgcagctcg	ccgaacaatt	caaacagaa	gcaatcgggt	tgctgcaaac	agccaccaag	3960
	caagcggagg	ctgctgctcc	cggtgtggaa	tccaagtggc	ggaccatcga	agccttcttg	4020
	gcgaagcata	tggtgaattt	catcagcggg	atacaataat	tagcaggctt	gtccactctg	4080
	cctggcaacc	ccgcgatagc	atcactgatg	gcattcacag	cctctatcac	cagcccgttc	4140
10	accacccaac	ataccctcct	gtttaacatc	ctggggggat	gggtggccgc	ccaacttgct	4200
	cctcccagcg	ctgcttctgc	tttcgtaggc	gccggcatcg	ctggagcggc	tggtggcagc	4260
	ataggccttg	ggaaggtgct	tggtgatatt	ttggcagggt	atggagcagg	gggtggcaggc	4320
	gcgctcgtag	cctttaaggt	catgagcggc	gagttgcctt	ccaccgagg	cctggttaac	4380
	ctactccctg	ctatctcttc	ccctggcgcc	ctagtcgtcg	gggtcgtgtg	cgcagcgata	4440
	ctgcgtcggc	acgtggggcc	aggggagggg	gctgtgcagt	ggatgaaccg	gctgatagcg	4500
	ttcgcttcgc	gggtaaacca	cgctctcccc	acgcactatg	tgcttgagag	cgcagctgca	4560
15	gcacgagtca	ctcagatcct	ctctagtctt	accatcactc	agctgctgaa	gaggcttcac	4620
	cagtggtatca	acgaggactg	ctccacgcca	tgctccggct	cggtggctaag	agatgttttg	4680
	gattggatat	gcacgggtgt	gactgatttc	aagacctggc	tccagtccaa	gctcctgcgc	4740
	cgattgcccgg	gagtcctcct	cttctcatgt	caacgtgggt	acaagggagt	ctggcggggc	4800
	gacggcatca	tgcaaaccaac	ctgccatgt	ggagcacaga	tcaccgggaca	tgtaaaaaac	4860
	ggttccatga	ggatcgtggg	gcctagggacc	tgtagtaaca	cgtggcatgg	aaatttcccc	4920
20	attaacgcgt	acaccacggg	cccctgcacg	ccctcccccg	cgccaaatta	ttctagggcg	4980
	ctgtggcggg	taggtgctga	ggagtacgtg	gaggttacgc	gggtggggga	tttccactac	5040
	gtgacgggca	tgaccactga	caacgtaaa	tgcccggtgc	aggttccggc	ccccgaattc	5100
	ttcacagaag	tggtgggggt	gcggttgca	aggtacgctc	cagcgtgcaa	accctctcta	5160
	cgggagggag	tcacatctct	ggtcgggctc	aatcaatacc	tggttgggtc	acagctccca	5220
	tgcgagcctg	aaccggtatg	agcagtgttc	acttccatgc	tcaccgaccc	ctccacattc	5280
25	acggcggaga	cggctaagcg	taggtcggcc	aggggatctc	ccccccctt	ggccagctca	5340
	tcagctagcc	agctgtctgc	gccttctctg	aaggcaacat	gcaactaccg	tcagtactcc	5400
	ccggacacct	acctcatcga	ggccaacctc	ctgtggcggc	aggagatggg	cgggaacatc	5460
	acccgcgtgg	agtcagaaaa	taagtagta	atthtggact	ctttcgagcc	gctccaaagc	5520
	gaggaggatg	agaggggagt	atccgttccg	gcggagatcc	tgccggaggc	caggaaatcc	5580
	cctcgagcga	tgcccatatg	ggcacgctcc	gattacaacc	ctccactgtt	agagtcctgg	5640
30	aaggaccggg	actacgtccc	tccagtggta	cacgggtgtc	cattgccgcc	tgccaaagcc	5700
	cctccgatac	caacttcacg	gaggaagagg	acggttgtcc	tgtcagaatc	taccgtgtct	5760
	tctgccttgg	cggagctcgc	cacagagacc	ttcggcagct	ccgaatcgtc	ggccgtcgac	5820
	agcggacagg	caacggcctc	tcctgaccag	ccctccgacg	acggcgacgc	gggatccgac	5880
	gttgagtctg	actcctccat	ccccccctt	gagggggagc	cgggggatcc	cgtatccagc	5940
	gacgggtctt	ggtctaccgt	aagcaggagg	gctagtggag	acgtcgtctg	ctgctcgatg	6000
	tcctacacct	ggacagggcg	cctgatcacg	ccatgcgctg	cggaggaaac	caagctgccc	6060
	atcaatgcac	tgagcaactc	ttgtctcgt	caccacaact	tggtctatgc	tacaacattc	6120
35	cgcagcgcaa	acctgcggca	gaagaaggtc	acctttgaca	gactgcagg	cctggacgac	6180
	cactaccggg	acgtgtctaa	ggagatgaag	gcgaaggcgt	ccacagttaa	ggctaaactt	6240
	ctatccgttg	aggaaacctg	taagctgacg	ccccacatt	cggccagatc	taaattttgg	6300
	tatggggcaa	aggacgtccg	gaacctatcc	agcaaggccg	ttaacacat	ccgctccgtg	6360
	tggaaggact	tgctggaaga	cactgagaca	ccaattgaca	ccaccatcat	ggcaaaaaat	6420
	gagggtttct	gcgtccaacc	agagaagggg	ggccgcaagc	cagctcgcct	tatcgtattc	6480
40	ccagatttgg	gggttcgtgt	gtgcgagaaa	atggcccttt	acgatgtgg	ctccaccctc	6540
	cctcaggccg	tgatgggtct	ttcatacggg	ttccaaact	ctcctggaca	gcgggtcgag	6600
	ttcctgggtg	atgcctggaa	agcgaagaaa	tgccctatgg	gcttcgcata	tgacacccgc	6660
	tggtttgact	caacggtcac	tgagaatgac	atccgtgttg	aggagtcaat	ctaccaatgt	6720
	tgtgactttg	ccccgaagc	cagacaggcc	ataaggctgc	tcacagagcg	gctttacatc	6780
	ggggggcccc	tgactaattc	taaaaggcag	aactgcggct	atcgccgggtg	cccgcgagc	6840
45	ggtgtactga	cgaacagctg	cggtaatacc	ctcacatggt	acttgaagge	cgtgcgggc	6900
	tgtcgagctg	cgaagctcca	ggactgcacg	atgctcgtat	gcggagacga	ccttgtcgtt	6960
	atctgtgaaa	gcgcggggag	ccaagaggac	gaggcgagcc	tacgggcctt	cacggaggct	7020
	atgactagat	actctgcccc	ccctggggac	ccgcccacac	cagaatacga	cttggagttg	7080
	ataacatcat	gctcctccaa	tggtgtcagtc	gcgcacgatg	catctggcaa	aaggggtgtac	7140
	tatctcacc	gtgacccca	cacccccctt	gcgcgggctg	cgtgggagac	agctagacac	7200
	actccagtea	attcctggct	aggcaacatc	atcatgtatg	cgccacctt	gtgggcaagg	7260
50	atgatctctg	tgactcattt	cttctccatc	cttctagctc	aggaacaa	tgaaaaagcc	7320
	ctgattgttc	agatctacgg	ggcctgttac	tccattgagc	cacttgacct	acctcagatc	7380
	attcaacgac	tccatggcct	tagcgcatct	tcactccata	gttactctcc	aggtgagatc	7440
	aataggggtg	cttcatgcct	caggaaactt	gggtaccg	ccttgcgagt	ctggagacat	7500

EP 1 043 399 A2

	cgggccagaa	gtgtccgcgc	taggctactg	tcccaggggg	ggagggctgc	cacttggtggc	7560
	aagtacctct	tcaactgggc	agtaaggacc	aagctcaaac	tcactccaat	cccggctgcg	7620
5	tcccagttgg	atztatccag	ctggttcggt	gctggttaca	gcgggggaga	catatatcac	7680
	agcctgtctc	gtgcccgacc	ccgctgggtc	atgtggtgcc	tactcctact	ttctgtaggg	7740
	gtaggcacat	atctactccc	caaccgatga	acggggagct	aaacactcca	ggccaatagg	7800
	ccatcctggt	tttttccctt	tttttttttc	tttttttttt	tttttttttt	tttttttttt	7860
	ttttctcctt	tttttttcct	cttttttttc	ttttctttcc	tttggtggct	ccatcttagc	7920
	cctagtcacg	gctagctgtg	aaaggccggt	gagccgcttg	actgcagaga	gtgctgatac	7980
10	tggcctctct	gcagatcaag	t				8001

15

20

25

30

35

40

45

50

55

EP 1 043 399 A2

<210> 11
 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus

5

<400> 11

10

15

20

25

30

35

40

45

50

55

```

gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggctcg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgccccc 240
gcgagactgc tagccagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcga gtgcccggg aggtctcgta gaccgtgcac catgagcacg aatccataaac 360
ctcaaaagaaa aaccaaaccgt aacaccaacg ggccgcgcac gatgaacaa gatggattgc 420
acgcagggtc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cagcagggc gttccttgcg cagctgtgct cgcgttcttc actgaagcgg 660
gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacatg ccatttcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgcgat caggatgac tggacgaaga gcatcagggg ctgcgcagag 900
ccgaactgtt cgcacggctc aagggcgcgca tgcgcgacgg cgaggatctc gtcgtgaccc 960
atggcgtatg ctgttgcgg aatatcatgg tggaaaatgg ccgctttctt ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgtgaaaga gcttggcggc gaatgggctg accgcttctc cgtgctttac ggtatcgccg 1140
ctcccgattc gcagcgatc gccttctatc gccttcttga cgaagtcttc tgagtttaa 1200
cagaccacaa cggtttccct ctagegggat caattccgac cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttttt 1320
tccaccata tgcgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctagggtctt tccccctctc gccaaaggaa tgcaaggctc gttgaatgtc 1440
tggaaggaa cagttctctt ggaagcttct tgaagcaaaa caacgtctgt agcgaccctt 1500
tgcaaggcag ggaacccccc acctggcgag aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc gccacaaacc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtaccccaat gtatgggatc tgatctgggg cctcgtgca catgctttac atgtgtttac 1740
tcgaggttaa aaaaagctta ggcctcccgga accacgggga cgtggttttc ctttgaataa 1800
cacgataata ccatgggcac gaatcctaaa cctcaagaa aaaccaaag taacaccaac 1860
cgccgccac aggacgtcaa gttcccgggc ggtggtcaga tgcgtcgttg agtttacctg 1920
ttgccgcgca ggggccccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggtcg 1980
caacctcgtg gaaggcgaca acctatcccc aaggctcgcc agcccgaggg tagggcctgg 2040
gctcagcccg ggtacccctg gccctctat ggcaatgagg gcttgggggtg ggcaggatgg 2100
ctcctgtcac ccggtggctc tgggectagt tggggcccca cggacccccc gcgtaggtcg 2160
cgcaatttgg gtaaggtcat cgataccctc acgtgcggct tcgccgatct catggggtac 2220
attcgcctcg tcggcgcccc cctagggggc gctgccaggg ccctggcgca tggcgctccg 2280
gttctggagg acggcgtgaa ctatgcaaca gggaatctgc ccggttgctc cttttctatc 2340
ttccttttgg ctttgcgtgc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400
gtatccggag tgtacctgt caccgaacgac tgctccaacg caagcattgt gtatgaggca 2460
gcggacatga tcatgcatac ccccggtgtc gtgacctgag ttcgggagaa caactcctcc 2520
cgctgctggg tagcgtcac tcccacgctc gcggccagga acgctagcgt cccactacg 2580
acgatacgac gccatgtcga ttgtctggt gggcggtctg ctctctgctc cgctatgtac 2640
gtgggagatc tctgcggatc tgttttcttc gtgcgccagc tgttcacctc ctgcctcgc 2700
cggcacgaga cagtacagga ctgcaattgc tcaatatatc ccggccacgt gacaggtcac 2760
cgatatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgcat 2820
ttactccgga tcccacaagc tgtcgtggat atggtggcgg gggcccatgg gggagtccca 2880
gcgggccttg cctactattc catggtgggg aactgggcta aggttctgat tgtgatgcta 2940
ctctttgccg gcgttgacgg gggaaacctat gtgacagggg ggacgatggc caaaaacacc 3000
ctcgggatta cgtccctctt ttaccccggt tcatcccgag aaatccagct tgtaaacacc 3060
aacggcagct ggcacatcaa caggactgcc ctgaaactgca atgactccct caacactggg 3120
ttccttgcgt cgctgttcta cgtgcacaag ttcaactcat ctggtatgcc agagcgcatg 3180
gccagctgca gccccatcga cgcgttcgct cagggtgggg ggcccatcac ttacaatgag 3240
tcacacagct cggaccagag gccttattgt tggcactacg caccocggcc gtgcggtatc 3300
gtaccccgcg cgacggtgtg tggctccagt tactgcttca ccccaagccc tgcgtggtg 3360
gggacgacgg accggttcgg cgtccctacg tacagtgggg gggagaatga gacggacgtg 3420
ctgcttctta acaacacgg gccgcgcaa ggcaactggg ttgctgttac atggatgaat 3480
agcactgggt tcaccaagac gtgcgggggc ccccgtgta acatcggggg gatcgccaat 3540

```

5

10

15

20

25

30

35

40

45

50

55

```

aaaaccttga cctgccccac ggactgcttc cggaagcacc ccgaggccac ttacaccaag 3600
tgtggttcgg ggcccttgggt gacaccaga tgcttggtcc actaccata caggctttgg 3660
cactaccctt gcaactgtcaa ctttaccatc ttcaagggtta ggatgtacgt ggggggagtg 3720
gagcacaggc tcgaagccgc atgcaattgg actcgaggag agcgtttgtaa cctggaggac 3780
agggacagat cagagcttag cccgctgctg ctgtctacaa cggagtggca ggtattgccc 3840
tgttccctca ccaccctacc ggctctgtcc actggtttga tccatctcca tcagaacgtc 3900
gtggacgtac aataacctgta cggatatagg tccgagggtg tctcctttgc aatcaaatgg 3960
gagtatgtcc tgttgctctt ccttctctcg ggggacggcg gcgtctgtgc ctgcttgttg 4020
atgatgtgct tgatagctca agctgaggcc gccctagaga acctgggtgt cctcaacgcg 4080
gcatccgtgg ccggggcgca tggcattctc tocttccctc tgttctctcg tgcctcctgg 4140
tacatcaagg gcaggctggt ccttggggcg gcatatgccc tctacggcgt atggccgcta 4200
ctcctgtccc tgctggcggt accaccacga gcatacgcca tggaccggga gatggcagca 4260
tcgtgcgagg gcgcgggttt cgttaggtctg atactcttga cctgttcacc gcactataag 4320
ctgttccctc ctaggctcct atggttggtta caatatttta tcaccagggc cgaggcacac 4380
ttgcaagtgt ggatccccc cctcaacggt cgggggggcc gcgatgccgt catcctccc 4440
acgtgcgcga tccaccacga gctaattctt accatcacca aaatcttgc cgcatactc 4500
gggtccactca tgggtgctcca ggctggtata accaaagtgc cgtacttctg gcgcgcacac 4560
ggggtcattc gtgcatgcat gctggtgctg aagggttgct ggggtcatta tgtccaaatg 4620
gctctcatga agttggccgc actgacaggt acgtacgttt atgaccatct caccctactg 4680
cgggactggg ccacgcggg cctacgagac ctgcggtg cagttagacc cgtcgtcttc 4740
tctgataagg agaccaaggt tatcacctgg ggggcagaca ccggtcgctg tggggacatc 4800
atccttggggc tgccccgtct cgccccgagg gggagggaga tacatctggg accggcagac 4860
agccttgaag ggcagggggt gcgactcttc gcgcctatta cggcctactc ccaacagacg 4920
cgaggcctac ttggctgcat catcactagc ctacagggcc gggacaggaa ccaaggtcga 4980
ggggagggtcc aagtgtctc caccgcaaca caatctttcc tggcgacctg cgtcaatggc 5040
gtgtgttggg ctgtctatca tgggtgcccgc tcaaagacct ttgcccggcc aaaggggccc 5100
atcacccaaa tgtacaccaa tgtggaccag gacctcgtcg gctggcaagc gccccccggg 5160
gcgcgttctc tgacaccatg cacctgcggc agctcggacc tttacttgggt caccgagcat 5220
gcccgtgtca ttccgggtgc ccggcggggc gacagcaggg ggaacctact ctccccagg 5280
cccgtctcct acttgaaggg ctcttcgggc ggtccactgc tctgcccctc ggggacagct 5340
gtgggcatct ttccgggtgc cgtgtgcacc cgaggggttg cgaaggcggt ggaactttga 5400
cccgtcgaag ctatgggaac cactatgcgg tcccgggtct tcacggacaa ctccgtccct 5460
ccggccgtac cgcagacatt ccagggtggc catctacacg cccctactgg tagcggcaag 5520
agcactaagg tgccggctgc gtatgcaggc caagggtata aggtgcttgt cctgaacccg 5580
tccgttcgcc ccacctagg tttcgggggc tatatgtcta aggcacatgg tatcgacct 5640
aacatcagaa tcgggggtaag gaccatcacc acgggtgccc ccatcacgta ctccacctat 5700
ggcaagtttc ttgccgacgg tgggtgtctct gggggcgccct atgacatcat aatatgtgat 5760
gagtgcacat caactgactc gaccactatc ctgggcatcg gcacagtcct ggaccaagcg 5820
gagaeggctg gagcgcgact cgtctgtctc gccaccgcta cgcctcggg atcgggtacc 5880
gtgccacatc caaacatcga ggaggtggct ctgtccagca ctggagaat ccccttttat 5940
ggcaaaagcca tccccatcga gaccatcaag ggggggaggc acctcatttt ctgccattcc 6000
aagaagaaat gtgatgagct cgccgcgaag ctgtccggcc tcggactcaa tctgttagca 6060
tattaccggg gccttgatgt atccgtcata ccaactagcg gagacgtcat tctcgtagca 6120
acggacgctc taatgacggg ctttaccggc gacttcgact cagtgatcga ctgcaatata 6180
tgtgtcacc ccacagtcga cttcagcctg gaccgcacct tcaccattga gacgacgacc 6240
gtgccaacaa acgcggtgtc acgctcgca cggcgaggca ggactggtag gggcaggatg 6300
ggcatttaca gggtttgtgac tccaggagaa cggccctcgg gcatgttcga ttctcgtgt 6360
ctgtgcgagt gctatgacgc gggctgtgct tggtaacgag tcacgcccgc cgagacctca 6420
gttaggttgc gggcttacct aaacacacca ggggtgccc tctgcccagg ccatctggag 6480
ttctgggaga gcgtctttac aggectcacc cacatagacg cccatttctt gtcccagact 6540
aagcaggcag gagacaactt cccctacctg gtagcatacc aggtacaggt gtgcgccagg 6600
gctcaggctc cacctccatc gtgggaccaa atgtggaagt gtctcatacg gctaagacct 6660
acgctgcacg ggccaacgcc cctgctgtat aggttgggag ccgttcaaaa cgaggttact 6720
accacacacc ccataaccaa atacatcatg gcatgcatgt cggctgacct ggaggtcgtc 6780
acgagcacct ggggtgctggt aggcggagtc ctagcagctc tggccgcgta ttgctgaca 6840
acaggcagcg tggtcattgt gggcaggatc atcttgtccg gaaagccggc catcattccc 6900
gacagggaag tcctttaccg ggagtctgat gagatggaag agtgccctc acacctccc 6960
tacatcgaac agggaaatgma caattcaaac agaaggcaat cgggttgtctg 7020
caaacagcca ccaagcaagc ggaggctgct gctcccgtgg tggaaatcaa gtggcggaac 7080
atcgaagcct tctggcgcaa gcatatgtgg aatttcatca gcgggataca atatttagca 7140
ggcttgtcca ctctgcctgg caaccccgcg atagcatcac tgatggcatt cacagctctc 7200
atcaccagcc cgctcaccac ccaacatacc ctctgttta acatcctggg gggatgggtg 7260
gcgcccacac ttgctcctcc cagcgcgtgt tctgctttcg taggcgccgg catcgctgga 7320
gcggctggtg gcagcatagg ccttggaag gtgcttgtgg atattttgg aggttatgga 7380
gcagggggtg caggcgcgct cgtggccttt aaggtcatga gcggcgagtt gccctcacc 7440
gaggacctgg ttaacctact cctgctatc ctctcccctg gcgccctagt cgtcggggtc 7500

```

	gtgtgcgcag	cgatactgcg	tccggcacgtg	ggcccagggg	aggggggctgt	gcagtgggatg	7560
	aaccggctga	tagcgttcgc	ttcgcgggggt	aaccacgtct	ccccacgca	ctatgtgcct	7620
	gagagcgacg	ctgcagcacg	agtcactcag	atcctctcta	gtcttaccat	cactcagctg	7680
5	ctgaagaggc	ttcaccagtg	gatcaacgag	gactgtctca	cgccatgctc	cggctcgtgg	7740
	ctaagagatg	tttgggattg	gatatgcacg	gtgttgactg	atttcaagac	ctggctccag	7800
	tccaagctcc	tgccgcgatt	gccgggagtc	cccttcttct	catgtcaacg	tgggtacaag	7860
	ggagtctggc	ggggcgacgg	catcatgcaa	accacctgcc	catgtggagc	acagatcacc	7920
	ggacatgtga	aaaacggttc	catgaggatc	gtggggccta	ggacctgtag	taacacgtgg	7980
	catggaacat	tccccattaa	cgcgtacacc	acgggccccct	gcacgccccct	cccggcgcca	8040
10	aattattcta	gggcgctgtg	gcgggtaggt	gctgaggagt	acgtggagggt	tacgcgggtg	8100
	ggggattttc	actacgtgac	gggcatgacc	actgacaacg	ttaaagtgcc	gtgtcagggt	8160
	ccggcccccg	aattcttcac	agaagtggat	ggggtgcggg	tgacagggtg	cgctccagcg	8220
	tgcaaaacccc	tccctacggga	ggaggtcaca	ttcctgggtcg	ggctcaatca	atacctgggt	8280
	gggtcacagc	tccccatgca	ggctgaaccg	gactgtagcag	tgctcacttc	catgtctacc	8340
	gacccctccc	acattacggc	ggagacggct	aagcgtaggc	tgccacgggg	atctcccccc	8400
15	cccttgggcca	gctcatcagc	tagccagctg	tctgcgcctt	ccttgaaggc	aacatgcact	8460
	acccgtcatg	actccccgga	cgtgacctc	atcgaggcca	acctcctgtg	gcggcaggag	8520
	atggggcgga	acatcacccc	cgtggagtca	gaaaataagg	tagtaatttt	ggactctttc	8580
	gagccgctcc	aagcggagga	ggatgagagg	ggagtatccg	ttccggcgga	gatcctgcgg	8640
	aggtccagga	aattccctcg	agcgatgcc	atatgggcac	gcccgatta	caaccctcca	8700
	ctgttagagt	cctggaagga	cccggactac	gtccctccag	tggtacacgg	gtgtccattg	8760
20	ccgcctgccca	aggccctcc	gataccacct	tcacggagga	agaggacggg	tgctcctgtc	8820
	gaatctaccg	tgtcttctgc	cttggcggag	ctcgccacag	agaccttcgg	cagctccgaa	8880
	tcgtcggccg	tcgacagcgg	cacggcaacg	gcctctcctg	accagccctc	cgacgacggc	8940
	gacgcgggat	ccgacgttga	gtcgtactcc	tccatgcccc	cccttgaggg	ggagccgggg	9000
	gatcccgatc	tcagcgacgg	gtcttggtct	accgtaagcg	aggaggctag	tgaggacgtc	9060
	gtctgtgctg	cgatgtccta	cacatggaca	ggcgccctga	tcacgccaatg	cgctgcggag	9120
25	gaaaccaagc	tgcccatcaa	tgcaactgagc	aactctttgc	tccgtcacca	caacttggtc	9180
	tatgctacaa	catctcgag	cgcaaacctg	cggcagaaga	aggctacctt	tgacagactg	9240
	cagggtcctgg	acgaccacta	ccgggacgtg	ctcaaggaga	tgaaggcgaa	ggcgtccaca	9300
	gttaaggcta	aacttctatc	cgtggaggaa	gcctgtaagc	tgacgcccc	acattcggcc	9360
	agatctaaat	ttggctatgg	ggcaaaaggac	gtccggaacc	tatccagcaa	ggccgttaac	9420
	cacatccgct	ccgtgtggaa	ggacttgctg	gaagacactg	agacaccaat	tgacaccacc	9480
30	atcatggcaa	aaaatgaggt	tttctgcgtc	caaccagaga	agggggggccg	caaggcagct	9540
	cgccctatcg	tattcccaaga	tttgggggtt	cgtgtgtgcg	agaaaatggc	cctttacgat	9600
	gtggctctcca	ccctccctca	ggccgtgatg	ggctcttcat	acggattcca	atacttccct	9660
	ggacagcggg	tcgagttcct	ggtgaatgcc	tgaaagcgca	agaaatgcc	tatgggcttc	9720
	gcatatgaca	cccgtgtttt	tgactcaacg	gtcactgaga	atgacatccg	tggttagggag	9780
35	tcaatctacc	aatgtttgta	cttggccccc	gaagccagac	aggccataag	gtcgtccaca	9840
	gagcggcttt	acatcggggg	ccccctgact	aattctaaag	ggcagaactg	cggctatcgc	9900
	cggtgccgcg	cgagcgggtg	actgacgacc	agctgcggta	ataccctcac	atgttacttg	9960
	aaggccgctg	cgccctgtcg	agctgcgaag	ctccaggact	gcacgatgct	cgtatgcgga	10020
	gacgaccttg	tcgttatctg	tgaaagcgcg	gggacccaag	aggacgaggc	gagcctacgg	10080
	gccttcacgg	aggctatgac	tagatactct	gccccccctg	gggacccgcc	caaaccagaa	10140
40	tacgacttgg	agttgataac	atcatgctcc	tccaatgtgt	cagtcgcgca	cgatgcacat	10200
	ggcaaaaagg	tgtactatct	cacccgtgac	cccaccaccc	cccttgccgcg	ggctgcgtgg	10260
	gagacagcta	gacacactcc	agtcaattcc	tggttaggca	acatcatcat	gtatgcgccc	10320
	accttgtggg	caaggatgat	cctgatgact	catttcttct	ccatccttct	agctcaggaa	10380
	caacttgaaa	aagccctaga	ttgtcagatc	tcaggggcct	gttactccat	tgagccactt	10440
	gacctacctc	agatcattca	acgactccat	ggccttagcg	cattttcact	ccatagtta	10500
45	tctccaggtg	agatcaatag	ggtggcttca	tgccctcagga	aacttggggg	accgcccttg	10560
	cgagtctgga	gacatcgggc	cagaagtgtc	cgcgctaggc	tactgtccca	gggggggagg	10620
	gctgccactt	gtggcaagta	cctcttcaac	tgggcagtaa	ggaccaagct	caaactcact	10680
	ccaatcccgg	ctgcgtccca	gttggattta	tccagctggg	tcgttgctgg	ttacagcggg	10740
	ggagacatat	atcacagcct	gtctcgtgcc	cgaccccgct	ggttcatgtg	gtgcctactc	10800
	ctactttctg	taggggtagg	catctatcta	ctccccaccc	gatgaaaggg	gagctaaaca	10860
50	ctccaggcca	ataggccatc	ctgttttttt	cccttttttt	ttttcttttt	tttttttttt	10920
	tttttttttt	tttttttttc	tccttttttt	ttcctctttt	tttccctttt	tttccctttg	10980
	tggtccatc	ttagccctag	tcacggctag	ctgtgaaagg	tccgtgagcc	gcttgactgc	11040
	agagagtgtc	gatactgggc	tctctgcaga	tcaagt			11076

Patentansprüche

1. Hepatitis C Virus (HCV) Zellkultursystem, das im wesentlichen eukaryontische Zellen umfaßt, die eingeschleustes HCV-spezifisches Genmaterial enthalten, dadurch gekennzeichnet,

daß die eukaryontischen Zellen humane Hepatomazellen sind und daß das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt ist, das die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zusätzlich ein selektierbares Markergen (Selektionsgen) umfaßt.

2. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet,

daß die Hepatomazellen von einer handelsüblichen Hepatomazelllinie abstammen.

3. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet,

daß die Hepatomazellen aus einer Hepatomaprimärzellkultur gewonnen sind.

4. HCV-RNA-Konstrukt, dadurch gekennzeichnet,

daß es die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zusätzlich ein selektierbares Markergen (Selektionsgen) umfaßt.

5. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet,

daß es eine Nukleotidsequenz gemäß einem der Sequenzprotokolle SEQ ID NO: 1 bis SEQ ID NO: 11 umfaßt.

6. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet,

daß die 3' NTR eine Nukleotidsequenz aufweist, die aus der Gruppe der nachfolgend aufgelisteten Nukleotidsequenzen (a) bis (i) ausgewählt ist:

(a) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTT
TTAGCTTTTTTTTTTTCTTTTTTTTGAGAGAGAGAGTCTCACTCTG
TTGCCCAGACTGGAGT

- 5 (b) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTCT TTTTTTTTCT TTTTTTTTGA GAGAGAGAGT
CTCACTCTGT TGCCAGACT GGAGC
- 10 (c) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTAATCTTT TTTTTTTTCT TTTTTTTTGA GAGAGAGAGT
CTCACTCTGT TGCCAGACT GCAGC
- 15 (d) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTCT TTTTTTTTTT TTTTTTTTTT TGAGAGAGAG
AGTCTCACTC TGTTGCCAG ACTGGAGT
- 20 (e) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTCT TTTTTTTTTT TTTTTTTTTT TGAGAGAGAG
AGTCTCACTC TGTTGCCAG ACTGGAGT
- 25 (f) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTCT TTTTTTTTTT TTTTTTTTTT TTGAGAGAGA
GAGTCTCACT CTGTTGCCCA GACTGGAGT
- 30 (g) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTCT TTTTTTTTTT CTTTTTTTTT GAGAGAGAGA
GTCTCACTCT GTTGCCAGAGA CTGGAGT
- 35 (h) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTTTAAT CTTTTTTTTT TTTTCCTTT TTTTGAGAGA
GAGAGTCTCA CTCTGTTGCC CAGACTGGAG T
- 40 (i) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTTAATC TTTTTTTTTT TTTTCTTTT TTTTGAGAG
AGAGAGTCTC ACTCTGTTGC CCAGACTGGA GT

- 45 7. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 6, dadurch gekennzeichnet,
daß das selektierbare Markergen ein Resistenzgen und insbesondere eine Antibiotikaresistenzgen ist.
- 50 8. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 6, dadurch gekennzeichnet,
daß das selektierbare Markergen ein Neomycinphosphotransferasegen ist.
- 55 9. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 8, dadurch gekennzeichnet,
daß das selektierbare Markergen strangabwärts der 5' NTR in die HCV-RNA integriert ist.
10. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 9, dadurch gekennzeichnet,

EP 1 043 399 A2

daß das selektierbare Markergen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA verbunden ist.

11. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 10, dadurch gekennzeichnet,

daß es ein integriertes Reportergen aufweist.

12. HCV-RNA-Konstrukt nach Anspruch 11, dadurch gekennzeichnet,

daß das Reportergen ein Gen aus der Gruppe der Luziferasogene, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasen), der GFP-Gene (green-fluorescence-protein-Gene), dem GUS-Gen (Glukuronidasen) und dem SEAP-Gen (Sezernierte-Alkalische-Phosphatase-Gen) ist.

13. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 11, dadurch gekennzeichnet,

daß deren Replikation die Expression eines (zellulären) Surrogatmarkergens beeinflusst.

14. HCV-RNA-Konstrukt nach einem der Ansprüche 11 bis 13, dadurch gekennzeichnet,

daß das Resistenzgen derart in das offene Leseraster der HCV-RNA inkliniert ist, daß es erst nach einer proteolytischen Prozessierung in eine aktive Form überführbar ist.

15. HCV-RNA-Konstrukt nach einem der Ansprüche 11 bis 14, dadurch gekennzeichnet,

daß das Reportergen und das selektierbare Markergen derart räumlich in dem Konstrukt angeordnet sind, daß sie gemeinsam ein Fusionsprotein exprimieren.

16. Zellkultursystem nach einem der Ansprüche 1 bis 3 dadurch gekennzeichnet,

daß das HCV-RNA-Konstrukt ein Konstrukt gemäß wenigstens einem der Ansprüche 4 bis 15 ist.

17. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet,

daß die das HCV-RNA-Konstrukt enthaltenden Zellen bei der DSMZ, Braunschweig, BRD, unter der Hinterlegungsnummer DSM ACC2394 (Laborbezeichnung HuBI 9-13) hinterlegt sind.

18. Verwendung eines Zellkultursystems nach einem der Ansprüche 1 bis 3 oder 16 bis 17 und/oder eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 15 zur Herstellung und/oder Evaluierung und/oder Testung von Therapeutika und/oder Diagnostika zur Behandlung von insbesondere HCV-Infektionen.

19. Verwendung eines Zellkultursystems nach einem der Ansprüche 1 bis 3 oder 16 bis 17 und/oder eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 15 zur Herstellung eines Impfstoffes gegen HCV-Infektionen.

20. Verwendung eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 15 zur Herstellung einer leberzell-spezifischen Genfahre für die Gentherapie .

21. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 15, dadurch gekennzeichnet,

daß es ein integriertes Fremdgen aufweist und dazu geeignet ist, dieses Fremdgen in eine Zielzelle einzuschleusen, die zur Expression dieses Fremdgens geeignet ist.

22. Verfahren zur Gewinnung von zellkultur-adaptierten Mutanten eines HCV-RNA-Konstrukts gemäß einem der Ansprüche 4 bis 15, wobei die Mutanten gegenüber dem HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen, dadurch gekennzeichnet,

daß man ein Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt mit Selektionsgen nach einem der Ansprüche 4 bis 15 ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, daß man die gewachsenen Zellklone erntet, und daß man

aus diesen Zellklonen die HCV-RNA-Konstrukte oder Teile davon isoliert.

23. Verfahren nach Anspruch 22, dadurch gekennzeichnet,

5 daß man die isolierten HCV-RNA-Konstrukte wenigstens einmal erneut passagiert, nämlich in Zellen eines Zellkultursystems nach Anspruch 1 einschleust, das dabei erhaltene Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial das isolierte HCV-RNA-Konstrukt mit Selektionsgen ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, die gewachsenen Zellklone

10

24. Verfahren zur Herstellung von Mutanten eines HCV-Vollängengenoms oder eines HCV-Teilgenoms oder eines beliebigen HCV- Konstrukts mit im Vergleich zu dem ursprünglichen HCV-Vollängengenom oder -Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz, dadurch gekennzeichnet,

15 daß man mit einem Verfahren nach Anspruch 22 oder 23 eine zellkultur-adaptierte Mutante eines HCV-RNA-Konstrukts herstellt und isoliert,

daß man die Nukleotid- und Aminosäuresequenz dieser Mutante bestimmt und durch Vergleich mit der Nukleotid- und Aminosäuresequenz des ursprünglichen HCV-RNA-Konstrukts die Art, Anzahl und Positionen der Nukleotid- und Aminosäuremutationen bestimmt,

20 und daß man diese Mutationen entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten, in ein (isoliertes) HCV- Vollängengenom oder ein HCV- Teilgenom oder ein beliebiges HCV-RNA-Konstrukt einführt.

25. Zellkultur-adaptiertes HCV-RNA-Konstrukt mit hoher Replikationseffizienz, dadurch gekennzeichnet,

25

daß es durch Nukleotid- und/oder Aminosäure -Mutationen von einem HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 15 ableitbar ist und daß es mit einem Verfahren nach einem der Ansprüche 22 bis 24 erhältlich ist.

30 26. Zellkultur-adaptiertes HCV-RNA-Konstrukt nach Anspruch 25, dadurch gekennzeichnet,

daß es einen oder mehrere der nachfolgend aufgeführten Aminosäureaustausche aufweist, nämlich 1283 arg -> gly und/oder 1383 glu -> ala und/oder 1577 lys -> arg und/oder 1609 lys -> glu und/oder 1936 pro -> ser und/oder 2163 glu -> gly und/oder 2330 lys -> glu und/oder 2442 ile -> val aufweist.

35

27. Zellkultur-adaptiertes HCV-RNA-Konstrukt nach Anspruch 25 oder 26, dadurch gekennzeichnet,

daß es einen oder mehrere der in Tabelle 3 aufgeführten Nukleotid- und/oder Aminosäureaustausche aufweist, wobei Tabelle 3 Bestandteil dieses Anspruchs ist.

40

28. Zellkultur-adaptierte Mutanten eines HCV-RNA-Konstrukts oder eines HCV-Vollängengenoms mit im Vergleich zu dem ursprünglichen HCV-RNA-Konstrukt oder dem ursprünglichen HCV-Vollängengenom erhöhter Replikationseffizienz, dadurch gekennzeichnet,

45 daß sie mit einem Verfahren erhältlich ist, bei dem man in einem zellkultur-adaptierten HCV-RNA-Konstrukt nach Anspruch 24 durch Sequenzanalyse und Sequenzvergleich die Art und Anzahl der Mutationen bestimmt und diese Mutationen in ein HCV-RNA-Konstrukt, insbesondere in ein HCV-RNA-Konstrukt gemäß einem der Ansprüche 4 bis 15, oder in ein (isoliertes) HCV-RNA-Vollängengenom einführt, entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten.

50

29. Hepatitis C Viruspartikel oder virus-ähnliche Partikel dadurch gekennzeichnet,

daß sie mit einem Verfahren nach einem der Ansprüche 22-24 erhältlich sind.

55 30. Zellen, infiziert mit Hepatitis C Viruspartikeln oder virus-ähnliche Partikeln gemäß Anspruch 29.

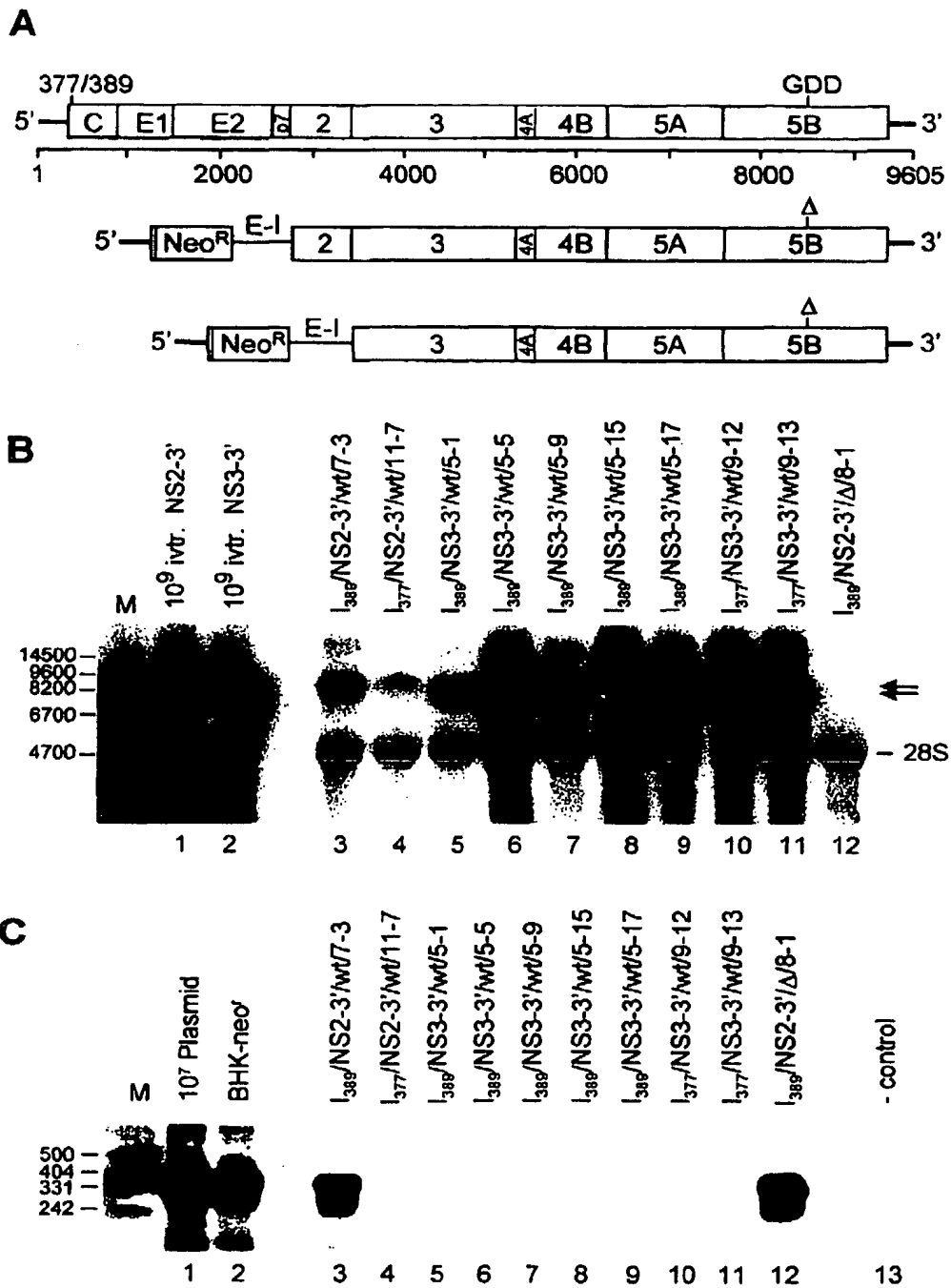


Fig. 1

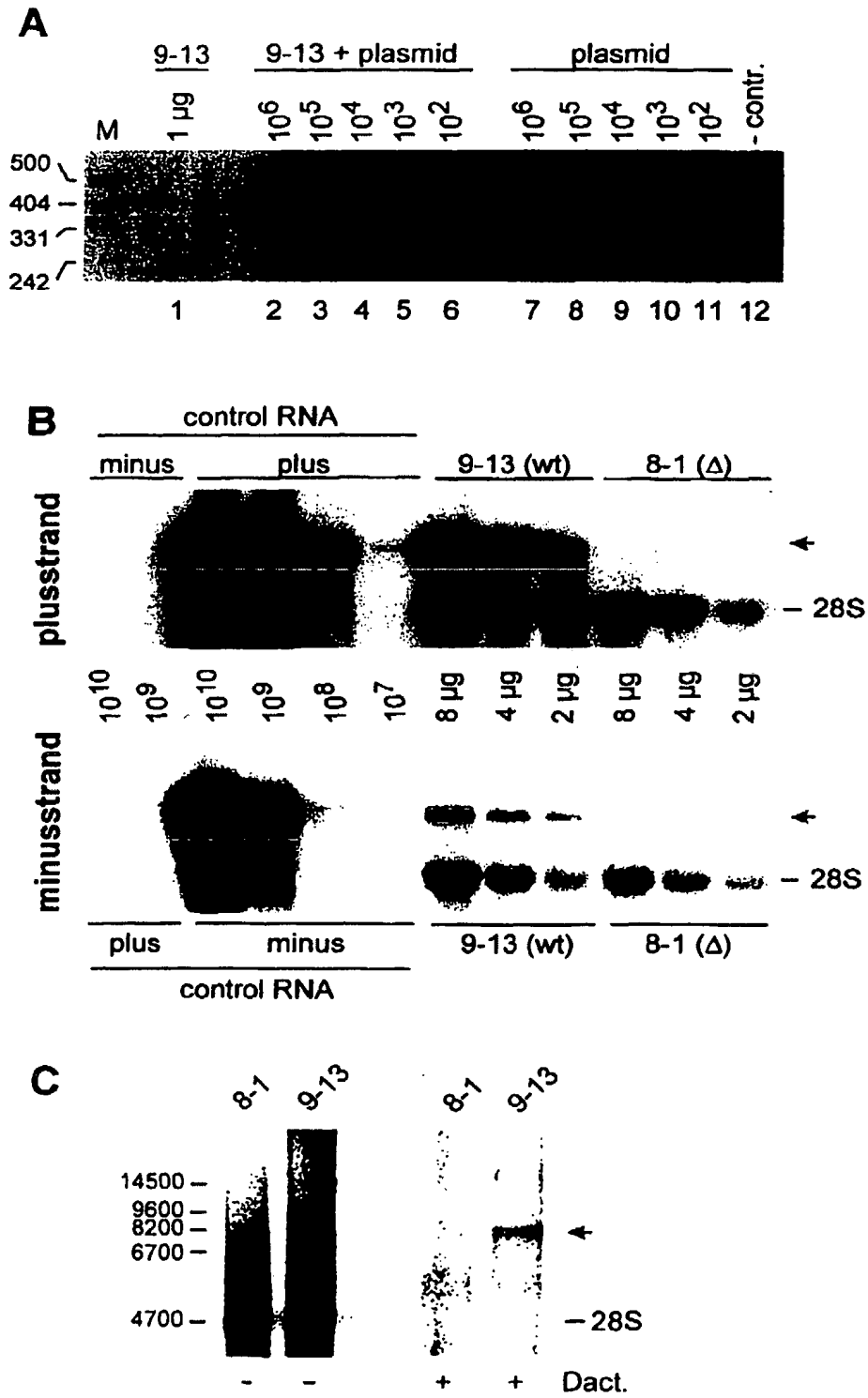


Fig. 2

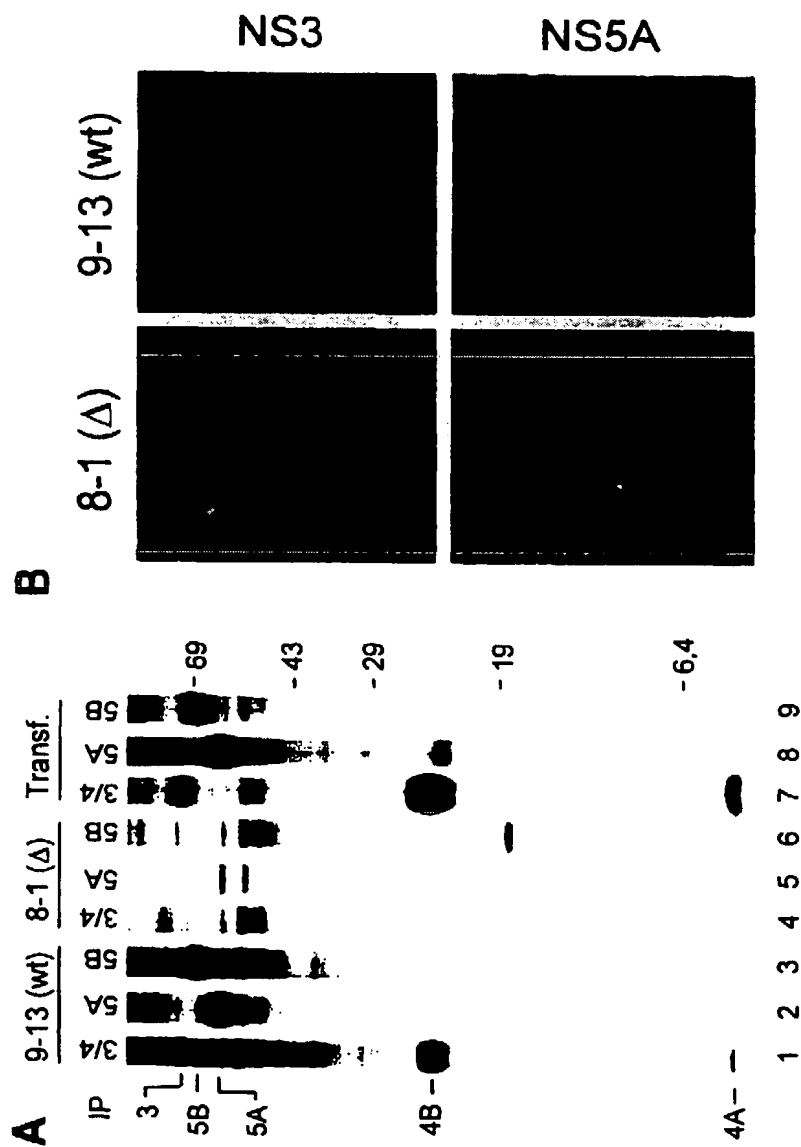


Fig. 3

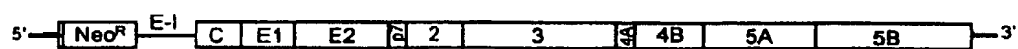


Fig. 4

A



B

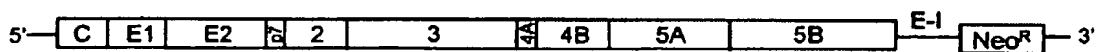
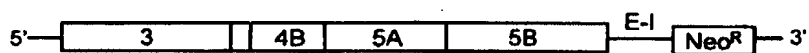


Fig. 5

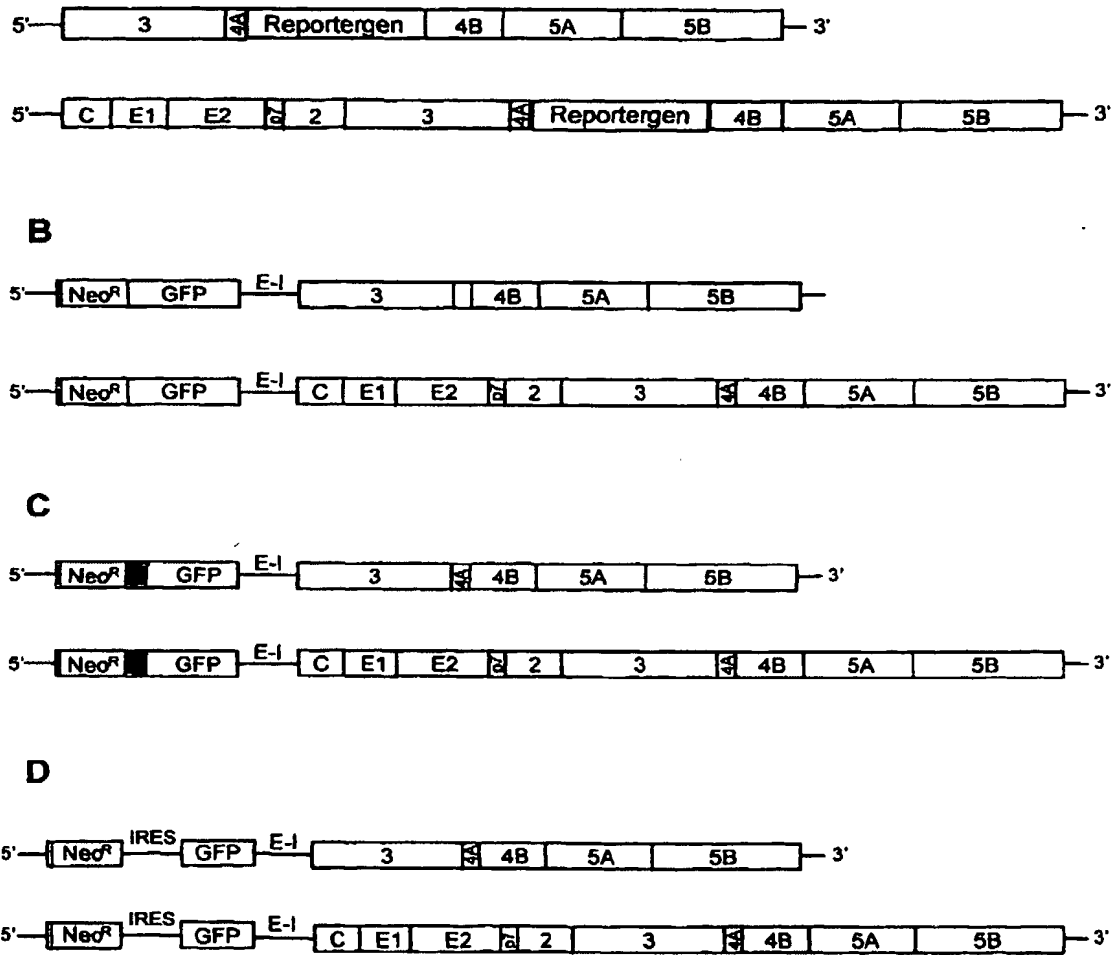


Fig. 6

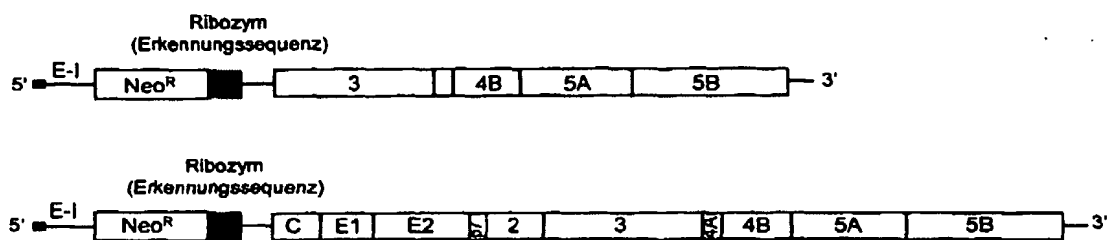


Fig. 7

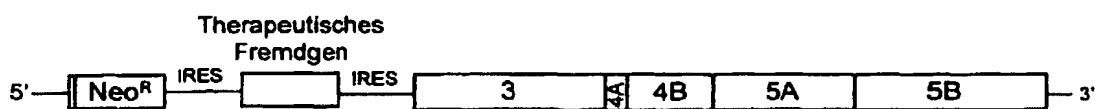


Fig. 8

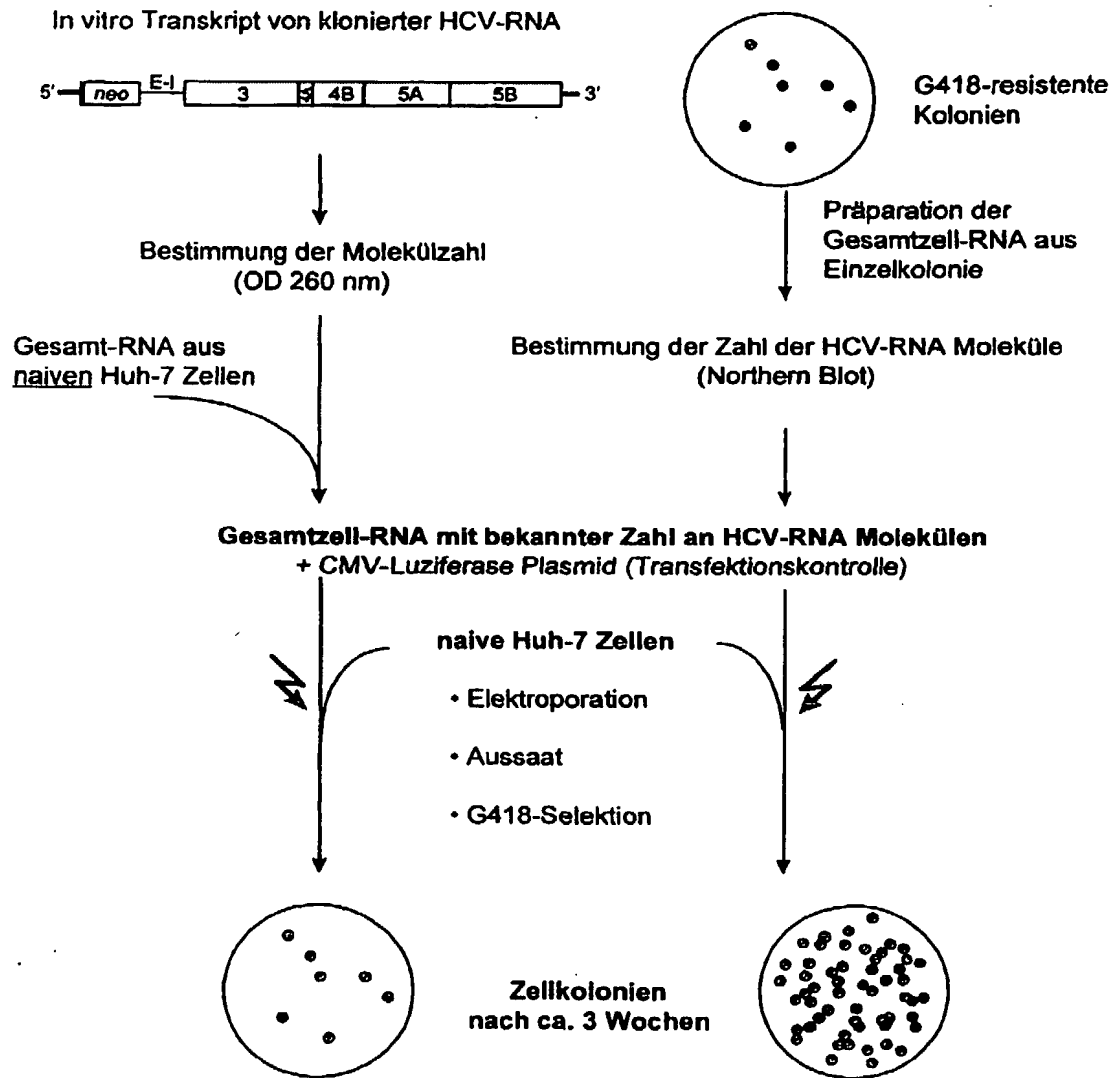


Fig. 9

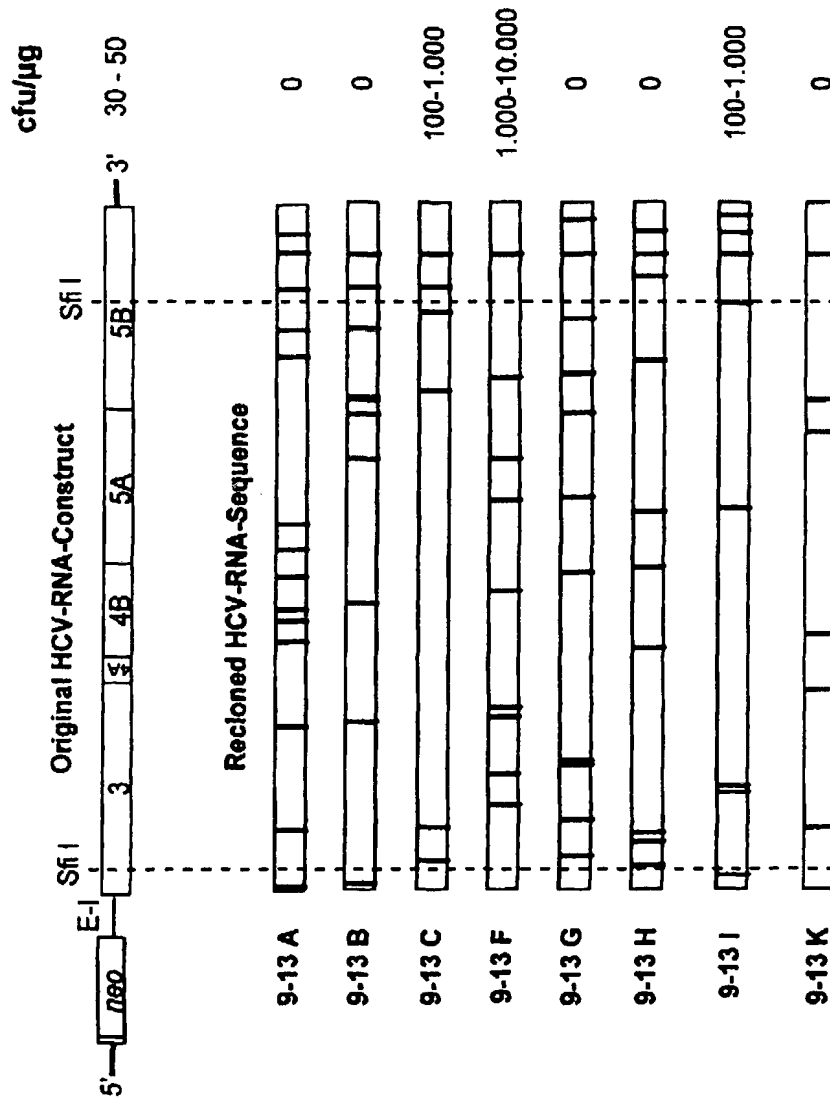


Fig. 10

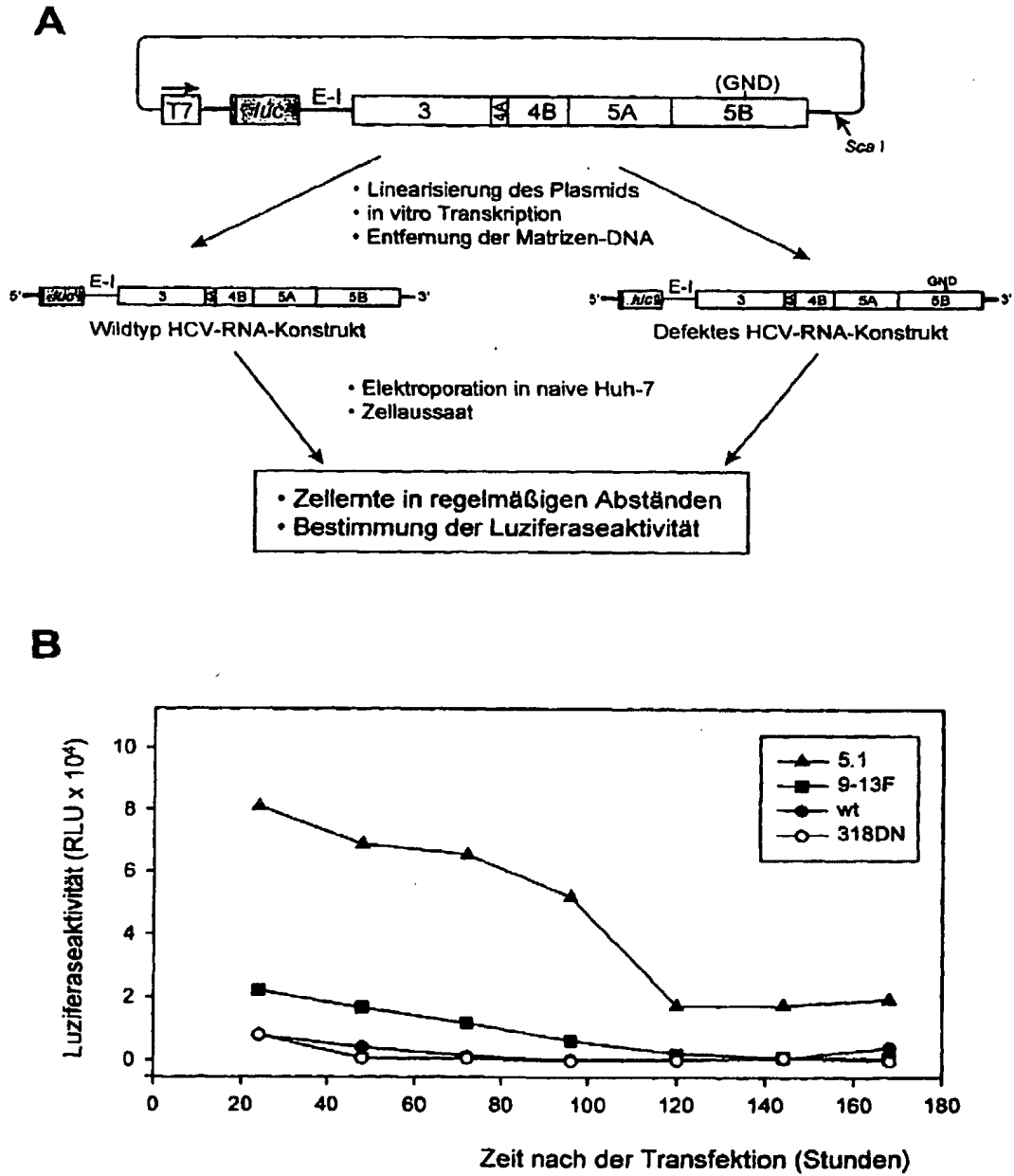


Fig. 11

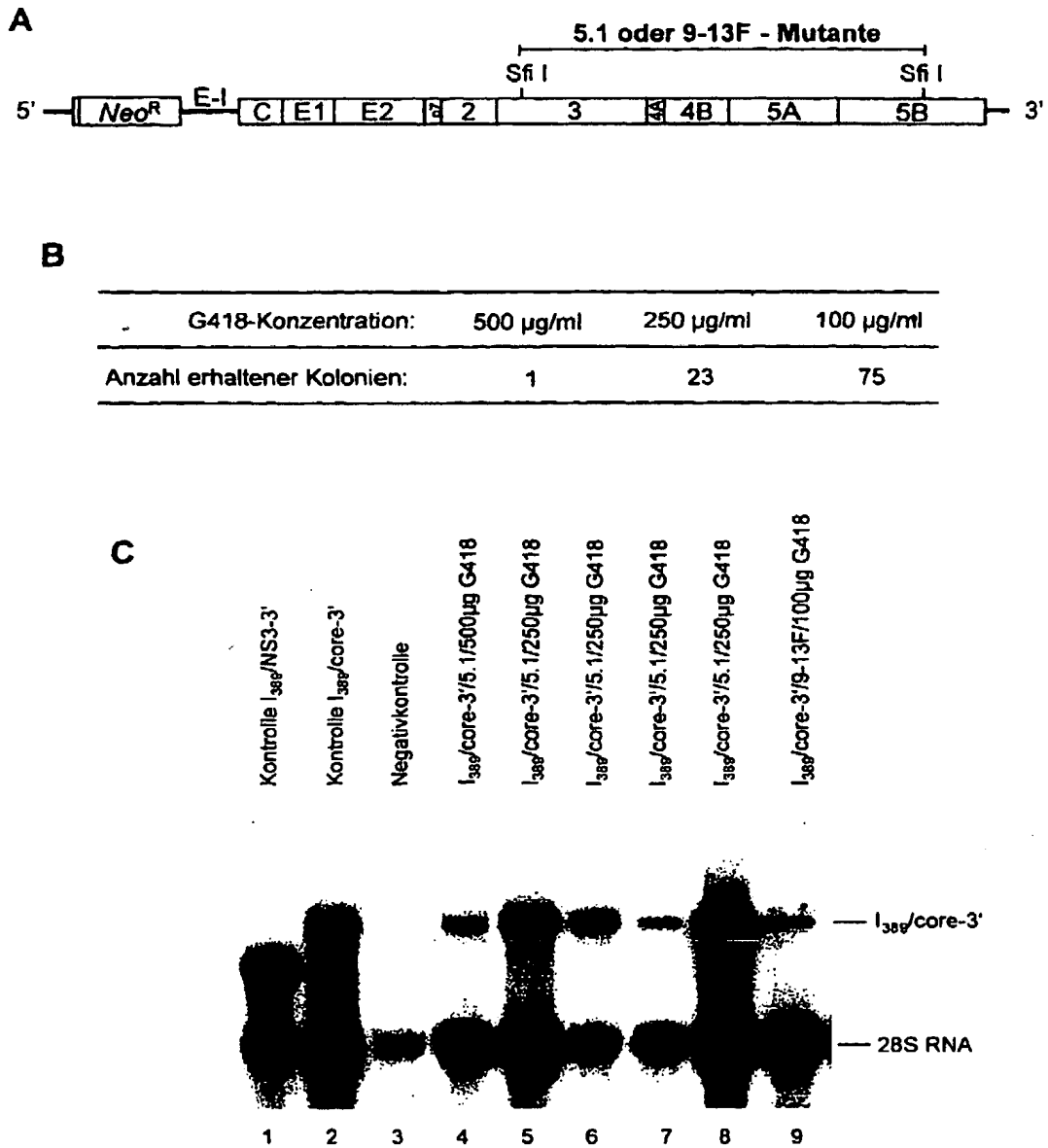


Fig. 12

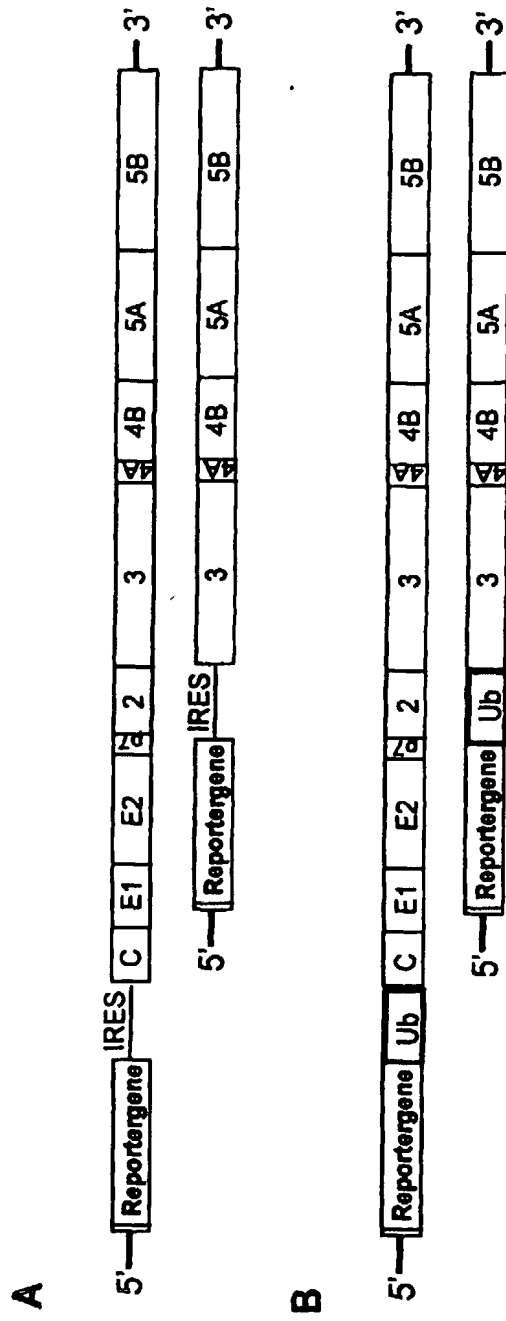


Fig. 13



Fig. 14

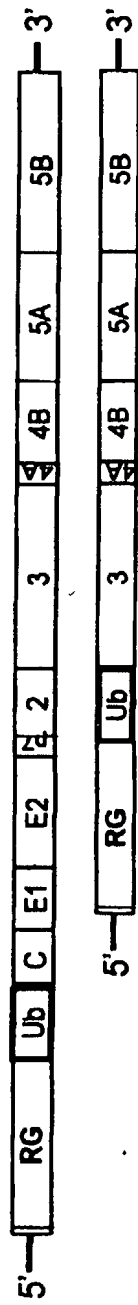


Fig. 15